



Deteção de Emoções em argumentos trocados no contexto de tomada de decisão em grupo

JOÃO TIAGO DA MOTA MOREIRA

Outubro de 2020

Deteção de Emoções em argumentos trocados no contexto de tomada de decisão em grupo

João Tiago Mota Moreira

**Dissertação para obtenção do Grau de Mestre em
Engenharia Informática, Área de Especialização em
Engenharia de Software**

**Orientador: Doutora Maria Goreti Carvalho Marreiros
Co-Orientador: Mestre Luís Manuel Silva Conceição**

Júri:

Presidente:

[Nome do Presidente, Categoria, Escola]

Vogais:

[Nome do Vogal1, Categoria, Escola]

Nome do Vogal2, Categoria, Escola

Resumo

O Homem é um ser senciente, pois experiencia, de forma consciente, vários sentimentos e várias emoções. As emoções estão presentes no seu dia a dia e na sua interação com o outro.

A Inteligência Artificial, é uma inteligência semelhante à humana exibida por mecanismos ou por software. Por essa razão, sempre houve a preocupação de entender o pensamento humano, para uma melhor “imitação”. O crescimento da Inteligência Artificial, levou ao aumento na pesquisa e no desenvolvimento de soluções sobre a detecção e a análise de emoções. Numa tomada de decisão o ser humano não utiliza apenas aspetos racionais para efetuar uma escolha, também utiliza as suas emoções. Um processo de tomada de decisão pode ser realizado em grupo ou individualmente. As tomadas de decisões em grupo serão focadas nesta dissertação. Para o auxílio no processo de tomada de decisão, vários autores investigaram e desenvolveram sistemas de apoio à decisão. Alguns desses sistemas consideram as emoções dos intervenientes no processo de tomada de decisão.

Nesta dissertação o enfoque está no estudo e na experimentação sobre a análise de emoções, bem como no impacto que os mesmos terão numa tomada de decisão. A análise de emoções é realizada em textos que traduzem processos de tomada de decisão.

O projeto desenvolveu-se no âmbito académico do Instituto Superior de Engenharia do Porto. E tem como objetivos: a detecção e a procura de emoções num determinado texto; e a verificação do seu impacto emocional no contexto do suporte de decisões de tomada de decisões em grupo.

Palavras-chave: Análise de sentimentos, Computação afetiva, sistemas de suporte de decisão, impacto emocional

Abstract

A human is a sentient being, because experience, in a conscious way, several sentiments and emotions. The emotions are present in their every day interactions with each other.

The Artificial Intelligence, is an intelligence similar to the human intelligence exhibited by mechanisms or by software. For that reason, always was a concern to understand the human thoughts, for a better "imitation". The growth of Artificial Intelligence, also made a raise of the research and development of solutions in the areas of emotion detection and analysis. In a decision making the human being does not apply just facts to make a choice, but also emotions. The decision making processes can be done in group or individually. The decisions that are made in group will be the focus of this dissertation. To help the process of decision making, several authors have researched and developed some systems to support decision making. Some of those support decision systems do not consider the member's emotions on the decision making process.

On this thesis the focus is on both the emotions studies and experimentation of emotion analysis, as well as on the impact of those emotions will have on a decision making process. The emotion analysis is done on texts which show decision making processes.

This project was developed in an academical environment in Instituto Superior de Engenharia do Porto. And have as goals: the detection and demand of emotions on a certain text and the verification of their emotional impact on the decision making process.

Agradecimentos

Agradeço, em primeiro lugar, aos meus orientadores, a Prof.^a Doutora Goreti Marreiros e o Me. Luís Conceição. A sua sabedoria, rigor, exigência e disponibilidade ao longo da realização deste trabalho, foram cruciais para a sua conclusão.

Aos meus pais e à minha irmã, pelo apoio e motivação.

À Eduarda e aos meus amigos agradeço as partilhas de conhecimento, os desabafos e a força que me deram, não só ao longo da realização deste trabalho, como também desde que nos conhecemos.

Conteúdo

Lista de Figuras	xi
Lista de Tabelas	xiii
Lista de Símbolos	xv
Lista de Acrónimos	xvii
1 Introdução	1
1.1 Enquadramento	1
1.2 Objetivos e Breve Descrição do Trabalho Realizado	2
1.3 Estrutura	2
2 Estado da arte	5
2.1 Análise de valor de negócio	5
2.1.1 Valor do negócio	7
2.1.2 Modelo de negócio	8
2.2 Processo de decisão	10
2.3 Processamento de linguagem natural	11
2.4 Computação afetiva	12
2.4.1 Análise de personalidade	12
<i>The big five</i>	13
Modelo PEN	15
2.4.2 Análise de emoções	15
O modelo OCC	16
Modelo emocional de <i>Ira Roseman</i>	17
<i>Hourglass of emotions</i>	17
2.4.3 Datasets e lexicons	19
Imdb dataset	20
Tweeter dataset	20
UN dataset	20
US debate dataset	20
WordNet	21
SenticNet	21
2.4.4 Abordagens utilizadas para determinar e classificar emoções	21
Agentes inteligentes	22
SVM	22
CRF	23
Naive Bayes classifier	23
Implementações	24
2.4.5 Impacto emocional em terceiros	28
Estudo de <i>Bersade</i>	28

Modelo EASI	28
Estudo de impacto de emoções em negociações de <i>Van Kleeef</i> . . .	29
2.5 Conclusões e discussões	30
3 Design e implementação da solução	31
3.1 Arquitetura do sistema	31
3.2 Alternativa da arquitetura do sistema	32
3.3 Desenvolvimento da solução	33
3.3.1 <i>TextAnalysisModule</i>	33
3.3.2 <i>SentimentModule</i>	35
4 Avaliação e experimentação	39
4.1 Métricas de avaliação e hipóteses	39
4.2 Metodologias de avaliação	40
4.2.1 Experiência com a detecção de emoções	40
4.2.2 Experiência com impacto emocional	41
4.3 Tipos de testes utilizados	45
5 Conclusões	47
5.1 Síntese do projeto	47
5.2 Limitações do projeto	48
5.3 Trabalhos futuros	48
Bibliografia	51

Lista de Figuras

2.1	Fuzy Front End adaptado de Belliveau, A. Griffin e Somermeyer (2004) . . .	6
2.2	Modelo de NCP adaptado de Belliveau, A. Griffin e Somermeyer (2004) . .	6
2.3	Modelo CANVAS	9
2.4	Modelo QFD	10
2.5	Modelo OCC adaptado de Ortony, Clore e Collins (1990)	16
2.6	Modelo <i>Hourglass of emotions</i> adaptado de Cambria, Livingstone et al. (2012)	19
2.7	Modelo de LST adaptado de Cambria, Poria et al. (2018)	27
3.1	Modelo de componentes do sistema	31
3.2	Modelo de componentes do módulo de análise emocional	32
3.3	Modelo alternativo de componentes do sistema	33
3.4	Exemplo de um possível dataset	34
4.1	Exemplo de um possível ficheiro resultado do algoritmo das colunas: Co- lumn1, IdDiscurso e Discurso	42
4.2	Exemplo de um possível ficheiro resultado do algoritmo das colunas: Interest, admiration, joy, surprise, sadness, disgust, fear, anger	42
4.3	Exemplo de um possível ficheiro resultado do algoritmo das colunas: Polarity e Sentics	43
4.4	Exemplo de um possível ficheiro resultado do algoritmo da coluna Impact .	44
4.5	Exemplo de um possível ficheiro resultado do algoritmo das colunas: E, Coe, Cm e Cd	44

Lista de Tabelas

2.1	Proposta longitudinal de valores.	8
2.2	Modelo <i>The big five</i> adaptado de P. J. Howard e J. M. Howard (1995) . .	14
2.3	Comparação entre abordagens adaptado de Pang e Lee (2004)	24
2.4	Comparação entre abordagens adaptado de Lin et al. (2006)	25
2.5	Resultados de teste <i>2-cross evaluation</i> adaptado de Agarwal et al. (2011) .	25
2.6	Resultados de teste <i>3 way evaluation</i> adaptado de Agarwal et al. (2011) . .	26
2.7	Comparação entre as precisões diferentes versões de adaptado de Cambria, Olsher e Rajagopal (2014)	26
2.8	Comportamentos de diversas emoções em ambientes cooperativos e competitivosG. van Kleef, C. De Dreu e A. Manstead (2010)	29
2.9	Comportamentos de diversas emoções em ambientes cooperativos e competitivos	30
3.1	Impacto de diversas emoções em ambientes de concordância e discordância	37
4.1	Resultados dos <i>F-scores</i> de vários módulos	41

Lista de Símbolos

Σ somatório

Lista de Acrónimos

API	<i>Application Programming Interface.</i>
CRF	<i>Conditional Random Fields.</i>
CRUD	Operações Create, Remove, Update, Delete.
csv	<i>Comma-separeted values.</i>
DAL	<i>Data Access Layer.</i>
FFM	<i>Five Factor Model.</i>
GECAD	Grupo de Investigação em Engenharia e Computação Inteligente para a Inovação e o Desenvolvimento.
QFD	<i>Quality Function Deployment.</i>
SADG	Sistemas de Apoio à Decisão em Grupo.
SVM	<i>Support vector machines.</i>

Capítulo 1

Introdução

O relatório foi desenvolvido no âmbito da unidade curricular tese/dissertação/estágio (TM-DEI) do mestrado em engenharia de informática do Instituto Superior de Engenharia do Porto, para documentar as etapas do desenvolvimento da mesma.

Neste capítulo será apresentado, de forma breve, o tema da dissertação: *Deteção de Emoções em argumentos trocados no contexto de tomada de decisão em grupo*. O capítulo será dividido em três partes. A primeira parte tratará do enquadramento, no qual é explicado e englobado o problema. A segunda parte tratará dos objetivos e de uma breve descrição do estudo. Por fim, a terceira parte tratará da estrutura da dissertação.

1.1 Enquadramento

Não obstante, as tomadas de decisões poderão não depender somente de um indivíduo, poderão depender de um pensamento coletivo. Numa empresa, realizam-se reuniões entre vários colaboradores para que uma decisão seja tomada. "Decision-making problems is the process of finding the best option from all of the feasible alternatives"(Chen 2000a, p. 1).

Numa decisão não estão apenas presentes factos ou argumentos. As emoções podem incitar uma deliberação. Elas aparecem regularmente e, por vezes, escondidas no processo da tomada da decisão. Podem surgir discreta ou indiscretamente nas palavras, nos gestos ou no tom de voz. "Across different types of decisions, important regularities appear in the underlying mechanisms through which emotions influence judgment and choice. Thus, emotion effects are neither random nor epiphenomenal" (Lerner 2004, p. 816). De modo a facilitar e melhorar as decisões tomadas, começaram a ser desenvolvidos sistemas de apoio à decisão.

Ao longo dos anos, começou a haver uma preocupação de conseguir detetar emoções, com o desenvolvimento da inteligência artificial e *machine learning*. Essa preocupação levou ao desenvolvimento de algoritmos que recebessem *tweets* ou *reviews* de filmes e detetassem as emoções. Daí, também resultou a deteção da polaridade dos mesmos (Agarwal et al. 2011; Barbosa e Feng 2010).

A temática desta dissertação consiste na análise de emoções presentes em texto, bem como, na determinação do impacto destas nas pessoas envolvidas em tomadas de decisões em grupo de problemas multicritério. Desta forma, o sistema de suporte de decisão é melhorado para agir conforme um ser humano.

A dissertação foi proposta pelo centro de investigação GECAD, Grupo de Investigação em Engenharia e Computação Inteligente para a Inovação e o Desenvolvimento. Neste centro de investigação já houve diversos trabalhos no campo de computação afetiva, de problemas multicritério e multiagentes.

1.2 Objetivos e Breve Descrição do Trabalho Realizado

O projeto que é exposto nesta dissertação é enquadrado na área de computação afetiva, no contexto de sistemas de apoio à decisão em grupo (SADG). O principal foco deste trabalho não são os SADG, mas a análise sentimental realizada sobre texto. No fundo, o que se pretende é um modelo que seja capaz de analisar e determinar as emoções presentes em texto. Em prol de atingir o objetivo principal, o trabalho foi dividido em duas partes ou dois objetivos:

- Detetar e identificar automaticamente emoções em mensagens trocadas pelos decisores em contexto de tomada de decisão em grupo;
- Determinar o impacto emocional de uma determinada afirmação nos restantes elementos participantes.

No entanto, para cumprir na totalidade os dois objetivos referidos ainda é necessário seguir as seguintes tarefas:

1. Realização do estado da arte sobre o suporte à tomada de decisão em grupo e computação afetiva;
2. Desenho e conceção de um modelo que permita detetar as emoções presentes em mensagens trocadas por decisores no contexto de decisão em grupo;
3. Desenho e conceção de um modelo que tenha a capacidade de prever o impacto emocional de uma determinada afirmação nos restantes decisores, num processo de tomada de decisão em grupo;
4. Elaboração dos testes necessários à validação do modelo;
5. Avaliação dos resultados obtidos nos testes.

Assim se pretende expor o trabalho realizado no âmbito desta dissertação.

1.3 Estrutura

A dissertação exposta foi desenvolvida de forma a que houvesse uma fácil compreensão por parte do leitor sobre a temática, o desenvolvimento e importância da mesma. A dissertação está também dividida em cinco capítulos.

No primeiro capítulo, a Introdução, foi dividido em três partes: o enquadramento, onde é exposto e explicado o problema da dissertação; os objetivos e uma breve descrição do trabalho realizado, onde são referidos os objetivos do trabalho no âmbito desta dissertação; e a estrutura.

Na primeira secção do primeiro capítulo (Estado da Arte), iremos analisar o valor do negócio e observar o estado da arte. Na última secção iremos aprofundar o conhecimento sobre

o contexto de tomada de decisões e da computação afetiva. No Estado da Arte serão abordados os seguintes pontos: a análise de personalidade e de emoções, *datasets* e *lexicons* e abordagens utilizadas para determinar e classificar emoções.

No terceiro capítulo o *design* e implementação da solução serão abordados. Ainda neste capítulo serão demonstradas e explicadas as escolhas realizadas ao longo do desenvolvimento da implementação e do próprio *design*.

Na experimentação e avaliação, o quarto capítulo, serão expostos os testes feitos ao longo do desenvolvimento da solução, bem como, a avaliação feita a esses mesmos testes.

O último capítulo consiste numa conclusão da dissertação. Neste capítulo serão identificados possíveis estudos a desenvolver no futuro, as limitações do estudo e as conclusões a que chegamos com o trabalho desenvolvido nesta dissertação.

Capítulo 2

Estado da arte

As emoções estão sempre presentes na tomada de decisões. Até mesmo, quando o indivíduo não tem consciência da implicação destas. "An integral component of human decision making is emotion, and this component could potentially be given to computers." (Picard 2000, p. 222). Por esta razão, no desenvolvimento de Sistemas de Apoio à Decisão em Grupo (SADG), é inevitável considerar a componente afetiva dos decisores, para que o resultado seja mais parecido com um ser humano.

Até aos dias de hoje, o estudo da análise de emoções ou computação afetiva tem grande impacto, sendo alvo de bastantes estudos. Damos como exemplos: as publicações do *tweeter* ou avaliações de filmes na IMDB.

Neste capítulo, inicialmente a análise do valor de negócio na 2.1. Depois iremos enumerar e descrever os estudos encontrados, na pesquisa, sobre o processos de tomada de decisão, 2.2 e de computação afetiva, 2.4 composta por: análise de personalidade (2.4.1), análise de emoções (2.4.2), *datasets e lexicons* (2.4.3), abordagens utilizadas (2.4.4) e estudos sobre o impacto emocional em terceiros (2.4.5). Por fim, na última secção (2.5) iremos fazer algumas considerações sobre a pesquisa.

2.1 Análise de valor de negócio

Com o aumento da procura das tecnologias por parte de empresas, dos próprios consumidores e da competitividade das empresas, criou-se uma elevada carência de inovação. Consequentemente, sentiu-se a necessidade de aumentar a pesquisa e de realizar mais estudos sobre a otimização do processo de desenvolvimento do *software*.

Na análise de valor, estudaremos a melhor maneira de aumentar o valor, de minimizar os custos da solução e da perda de qualidade do problema inserido na secção 1.1. Os pontos a serem abordados nesta secção serão:

- Indicar os métodos e técnicas adequadas para o tema;
- Indicar a proposta de valor;
- Analisar e integrar os itens *value for customer*, *percieved value* e *sacrifices*;
- Desenvolver um modelo de negócio.

Neste estudo, procuramos associar aos sistemas de apoio à tomada de decisão em grupo, (Sistemas de Apoio à Decisão em Grupo (SADG)) a análise de emoções subjacentes aos argumentos de cada indivíduo desse mesmo grupo. Assim, de certa forma, a tomada de

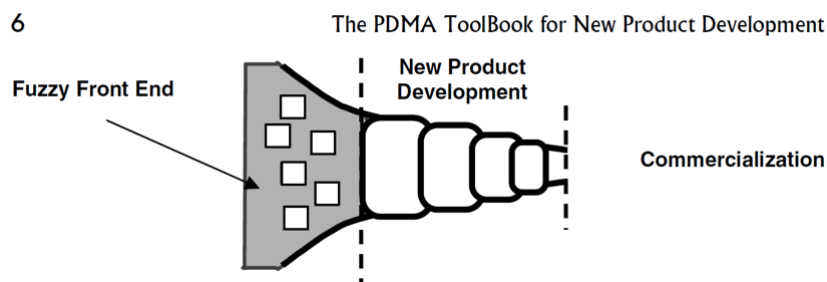


Figura 2.1: Fuzy Front End adaptado de Belliveau, A. Griffin e Somermeyer (2004)

decisão torna-se mais “humana”. Por outras palavras, tal como um ser humano é invadido pelas emoções numa tomada de decisão, também os SADG terão em conta as mesmas. Peter Koen desenvolveu o modelo *New Concept Development* (NCD), no qual o processo de inovação divide-se em três partes: a parte da frente da inovação (*Fuzzy Front End* (FFE)); o processo do desenvolvimento do produto novo (*the new product development* (NPD) *process*); e a comercialização (figura 2.1). Os autores referem ainda a importância do FFE no aumento da valorização e do sucesso. (Belliveau, A. Griffin e Somermeyer 2004).

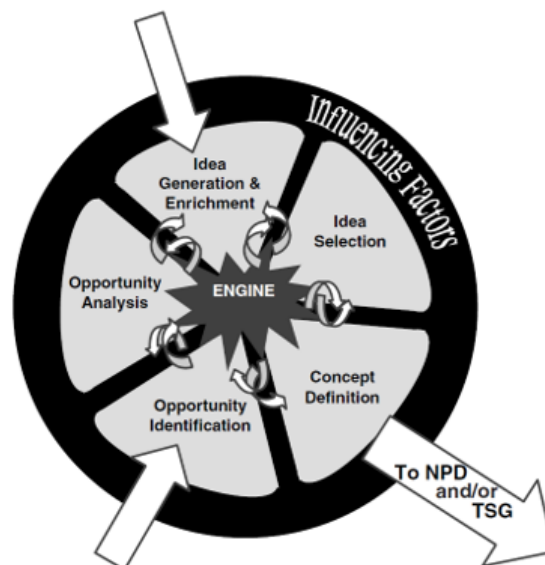


Figura 2.2: Modelo de NCP adaptado de Belliveau, A. Griffin e Somermeyer (2004)

O modelo NCP é dividido por três partes (figura 2.2):

- O motor (*engine*) é a liderança, cultura e a estratégia da organização, que controla os cinco fatores principais;
- A área interna que define os cinco fatores principais (identificação da oportunidade, análise da oportunidade, geração da ideia e enriquecimento da mesma, seleção da ideia e definição do conceito) do FFE;

- Os fatores que influenciam são: a capacidade da empresa, o mundo externo (canais de distribuição, leis, políticas do governo, clientes, competidores e clima político-económico) e as ciências (internas e externas), sendo fatores que não são controlados pela empresa.

Com o aumento de decisões em grupo nas grandes empresas e com o aumento da utilização de SADG, apareceram mais estudos. Assim como sobre o tema principal da dissertação, a análise de emoções. Porém, não há muitos estudos dentro deste contexto, como os apresentados na secção 2.4.2. Ou seja, existem estudos sobre o tema análise de emoções ou computação afetiva, mas não abordam o contexto dos sistemas de apoio à tomada de decisão em grupo (SADG). Este facto, apresenta uma oportunidade para o Grupo de Investigação em Engenharia e Computação Inteligente para a Inovação e o Desenvolvimento (GECAD).

Após a análise e identificação da oportunidade, executou-se uma fase de criação e enriquecimento da ideia que resultou na realização de um módulo que analisasse um determinado texto, determinasse as emoções presentes no mesmo e ainda o impacto provocado por elas. Na fase de seleção trata-se de obter a seleção dos melhores algoritmos para a solução do problema.

Em conclusão, o último elemento chave do NCD, é a fase de definição do conceito que consiste na implementação e teste do módulo da análise de emoções. O teste do módulo da análise de emoções será através da execução do módulo, utilizando um *dataset* como *input* de dados.

2.1.1 Valor do negócio

A criação de um valor é essencial para qualquer negócio, tal como expressou Woodall (2003). Porém, ainda fica por responder o que é o valor do potencial negócio. Como Walters e Lancaster (2000) afirmou, o valor de algo é um conceito que é difícil de obter, perceber, modelar e conceptualizar; certos autores consideram a criação do valor a diferença entre os benefícios e os sacrifícios. Como se pode concluir com o referido, anteriormente, é necessário fazer um estudo dos benefícios e sacrifícios que um cliente poderá ter ao utilizar um determinado produto ou serviço.

A proposta de valor, como já referida na secção 2.1, passa por associar aos sistemas de apoio à tomada de decisão em grupo (SADG), a análise de emoções subjacentes aos argumentos de cada individuo desse grupo.

Os potenciais clientes da área, ou seja, os utilizadores de SADG terão o valor detalhado na tabela 2.1, sob a forma de uma proposta de valores longitudinais. Os clientes serão beneficiados com esta ajuda, pois realizarão melhores decisões.

Os clientes terão quatro estados em relação à compra do serviço em questão, a pré-compra, a compra, pós-compra e pós-utilização. Na pré-compra o cliente terá os seguintes benefícios: um serviço que tome decisões considerando as suas emoções; outro serviço que otimize o seu tempo; e, ainda outro, que melhore as suas decisões. Na compra beneficiará da qualidade, da flexibilidade e da confiança. Pós-compra quer tempo economizado e que se exprima como esperado. Por fim, na após-utilização terá satisfação e melhores decisões tomadas.

Tabela 2.1: Proposta longitudinal de valores.

	Benefícios	Sacrifícios
Pré-Compra	Decisões considerando os seus sentimentos Otimize o seu tempo Melhore as suas decisões	Custo
Compra	Qualidade Flexibilidade Confiança	Custo
Pós-Compra	Tempo economizado Que se exprima como esperado	Tempo de Aprendizagem
Após-Utilização	Satisfação Melhores decisões tomadas	

2.1.2 Modelo de negócio

O modelo de negócio foi dividido em nove partes: parceiros-chave, atividades-chave, recursos-chave, proposta de valor, relações com o cliente, canais de distribuição, segmentação de clientes, estrutura de custos e fontes de receita.

Para a concretização desta dissertação, os parceiros-chave são empresas de *software* e centros de investigação e de desenvolvimento. A empresa de *software* auxiliaria a divulgar e a replicar protótipos dos algoritmos desenvolvidos. E os centros de investigação e de desenvolvimento deveriam fornecer recursos para todo o processo de investigação e desenvolvimento.

As atividades-chave, que serão executadas no decorrer desta dissertação, deverão desenvolver e investigar algoritmos de análise de emoções e de as integrar em sistemas de apoio às tomadas de decisões.

As equipas de desenvolvimento e investigação de um centro de investigação e os projetos de desenvolvimento são os recursos chave. As equipas referidas são o conjunto de pessoas que irão executar as atividades chave explicadas anteriormente.

A proposta de valor reside na determinação das emoções num texto, assim como o impacto emocional das mesmas, futuramente aplicáveis em Sistemas de apoio às tomadas de decisão, de modo a melhorar as decisões tomadas por esses sistemas.

Para reforçar as relações com os clientes, as empresas de *software*, prestarão suporte de utilizadores do *software*. Essas empresas também terão como trabalho um canal de distribuição e, ainda, publicações em revistas científicas. Deste modo pretende-se alcançar as administrações de grandes empresas, que são as que utilizam mais o SADG e a administração pública. Ambos constituem os grupos de possíveis clientes, que são o resultado de uma segmentação de clientes.

A estrutura de custos estará relacionada com os recursos humanos (equipas de investigação, desenvolvimento e suporte) e custos para as infraestruturas e licenças. E as principais receitas serão retiradas nas vendas de *software* produzido.

O modelo de negócio explicado nesta subsecção encontra-se representado sob a forma de um *Business Model Canvas* na figura 2.3.

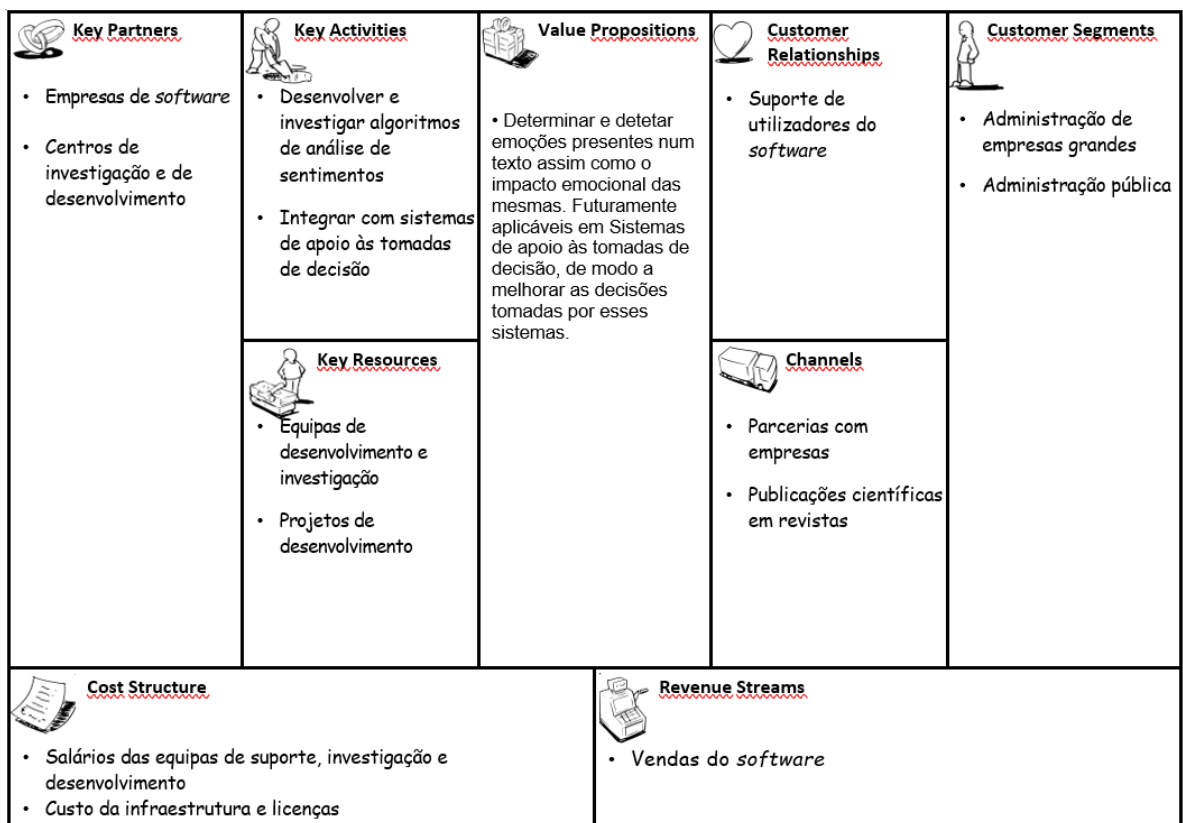


Figura 2.3: Modelo CANVAS

De modo a obter uma melhor qualidade de solução, de acordo com as necessidades dos clientes, foi utilizado um modelo *Quality Function Deployment* (QFD). O modelo QFD é um sistema que ajuda no processo de desenvolvimento de um produto ou serviço. No caso desta dissertação, o QFD ajudará no desenvolvimento de um serviço. O processo deste sistema divide-se em quatro passos:

- Requisitos do cliente - O que o cliente ou mercado espera do serviço;
- Funcionalidades da solução - O que a solução vai apresentar;
- Relevância para o cliente da e cada solução - É um valor numérico que se refere à importância de um requisito para o cliente;
- Pesagem da importância - Valor numérico que refere a importância da funcionalidade da solução para o cliente.

Na figura 2.4 está representado o modelo *Quality Function Deployment* (QFD) desenvolvido para esta dissertação.

Os requisitos do cliente passam pela deteção automática das emoções em mensagens. Este tipo de requisito tem um valor numérico (1) de relevância para o cliente, visto que não vai ser esse requisito a fornecer dados diretamente ao cliente. Porém, não deixa de ser importante. A identificação das emoções presentes nas mensagens e a previsão do impacto emocional de uma determinada afirmação, tem 2 valores numéricos de relevância para o cliente.

As funcionalidades da solução deste estudo são: módulo de tratamento de texto; módulo de análise de emoções; módulo de calcular o impacto emocional e o módulo de resultados. Para além da definição dos requisitos do cliente e das funcionalidades da solução, também foram determinadas as relações entre funcionalidades com requisitos e funcionalidades com funcionalidades. A pesagem da importância foi calculada pela multiplicação do valor da relevância, para o cliente, com o grau de ligação entre todos os requisitos que têm relação com a funcionalidade em questão.

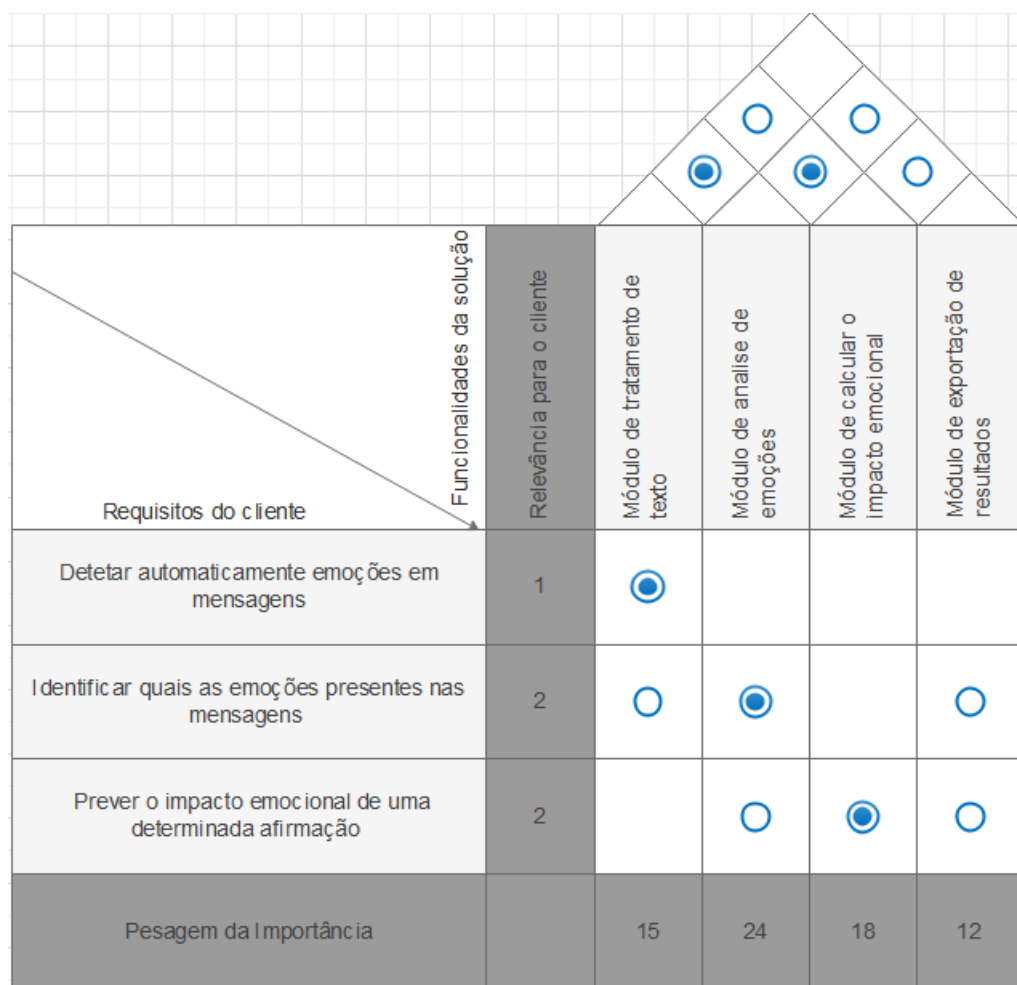


Figura 2.4: Modelo QFD

2.2 Processo de decisão

Saaty (2008, p. 83) disse, "We are all fundamentally decision makers". A citação anterior frisa a importância das tomadas de decisões no ser humano, visto que realizamos, diariamente, tomadas de decisões em todas as nossas ações, (Saaty e Peniwati 2013). Uma decisão é a escolha de uma ou várias soluções face a um problema, (Chen 2000b).

Na atualidade, a maior parte das decisões na área de gestão são tomadas em grupo, (F. Luthans, B. C. Luthans e K. W. Luthans 2015). Esta constatação resultou num aumento do estudo do tema de tomada de decisões em grupo, sendo o modo preferido para tomar decisões das organizações, (Moon et al. 2003). O estudo do processo de decisão é feito com o objetivo de tornar a tomada de decisão mais eficaz, ou seja, decisões mais corretas em menos tempo. O processo de tomada de decisão em grupo é executado entre várias pessoas. Juntos avaliam e analisam um conjunto de variáveis para determinar a solução mais acertada para a maior parte das pessoas em questão. As pessoas envolvidas podem estar todas juntas num só local ou dispersas geograficamente (F. Luthans, B. C. Luthans e K. W. Luthans 2015; Palomares, Martínez e Herrera 2014).

Ao longo dos anos, vários investigadores apresentaram diferentes opiniões sobre a melhor maneira de tomar uma decisão. Debatem se a melhor maneira será em grupo (Hilgard 1953) ou individualmente, (Buehler, Messervey e D. Griffin 2005). Buehler, Messervey e D. Griffin (2005) afirmaram que nas decisões em grupo, as pessoas intervenientes correm mais riscos. E demonstraram que a maior parte do tempo dos decisores não é gasto na tomada de decisões, mas em questões sociais. Dependendo do número de intervenientes no processo de tomada de decisão e do próprio processo, a tomada de decisão em grupo pode ser melhor do que a individual (Moon et al. 2003).

Pelo que foi referido, a tomada de decisões em grupo permite que se tomem decisões de qualidade, fazendo com que haja mais ajuda dos pares e com que se possa integrar mais facilmente pessoas nas decisões (Kaner 2014). Estes fatores tornam a tomada de decisão em grupo, muito apelativa para empresas. Devido à globalização, surgiram problemas da distância física e temporal. Ou seja, os integrantes dos processos de tomada de decisão estavam distantes entre si e muitas vezes em fusos horários diferentes. Para solucionar esses problemas surgiram as teleconferências, que permitiam que pessoas de diversos pontos geográficos pudessem realizar reuniões.

Nesta secção introduziu-se o tema da tomada de decisão e dos seus diferentes processos, que faz parte do tema da dissertação.

2.3 Processamento de linguagem natural

Nesta secção tentaremos explicar o que é o processamento de linguagem natural, bem como a ferramenta mais utilizada para a executar.

O processamento de linguagem natural, ou *Natural Language processing* (NLP), é o processamento ou análise das linguagens naturais. "By "natural language" we mean a language that is used for everyday communication by humans; languages such as English, Hindi, or Portuguese" (Bird, Klein e Loper 2009, p. ix). O processamento da linguagem natural é o estudo de formas utilizado pelo sistema computacional para analisar e compreender as linguagens não computacionais.

Este processamento é um pouco complexo, por ser necessária a criação de estruturas de dados que representem uma espécie de dicionários. Essas estruturas definirão as palavras para serem utilizadas, posteriormente, na análise de frases.

O processamento de linguagem natural estende-se também à deteção de fala por áudio, o que também permite a um sistema computacional detetar as palavras presentes nesse

mesmo áudio. Porém, esta área do processamento de linguagem natural não será relevante para esta dissertação.

Atualmente, existem várias ferramentas que oferecem um conjunto de algoritmos que processam a linguagem natural. A principal ferramenta vai ser abordada e explicada, a *NLTK*¹.

A *Natural Language Tool Kit*, ou NLTK, é uma das ferramentas mais utilizadas para processamento de linguagem natural. A ferramenta é compatível apenas em programas de linguagem *Python* (Loper e Bird 2002).

A NLTK é composta por diversos módulos independentes, cada um com a sua funcionalidade: como imprimir uma árvore de divisões, um módulo de testes ou de gerar texto. Obedecendo ao contexto desta dissertação, iremos explicar dos módulos *tokenizer* e apresentar o *corpus*.

O módulo *tokenizer* é utilizado para dividir um texto em palavras ou em frases. A forma como este módulo divide o texto em frases passa por dois processos. Primeiro, pela divisão do texto em *tokens*, que são um conjunto de caracteres separados por um espaço. Depois verifica o seu tipo. Isto, para evitar que o algoritmo divida as frases apenas por pontuação, visto que pode haver abreviações de palavras, como por exemplo "Dr."

O módulo *corpus* é o módulo que permite ler ficheiros em diversos formatos e permite ler ficheiros externos aos ficheiros internos da biblioteca.

2.4 Computação afetiva

A computação afetiva é a denominação de um ramo em Inteligência Artificial que estuda a parte afetiva de um ser humano e a tenta replicar em computadores. Esta área engloba disciplinas como a neurociência, psicologia e ciências da computação. O propósito da computação afetiva é detetar alterações na afetividade do utilizador e a perceção e motivação do mesmo (Serbedzija e Fairclough 2009).

Como Picard (2000) afirmou, o papel de emoções na afetividade e perceção humana é essencial. Assim, o sistema computacional, capaz de analisar os afetos, irá ajudar o próprio a tomar melhores decisões e consequentemente melhorará as decisões tomadas pelo ser humano. Factos que provam que as emoções são importantes num processo de tomada de decisão.

Nesta secção iremos abordar diferentes tópicos relevantes para a computação afetiva: a definição de uma pessoa, ou seja, a sua personalidade; as emoções; as abordagens para determinar e classificar emoções; e impacto emocional em terceiros. Nesses quatro tópicos vão ser explicados os modelos mais utilizados e como são aplicados no contexto computacional.

2.4.1 Análise de personalidade

A personalidade de alguém ajuda a obter um possível padrão de um indivíduo. É sempre importante conhecer este fator, pois pode ser vantajoso no processo de tomada de decisão (Carneiro 2018). Embora a análise de personalidade não seja muito utilizada no contexto desta dissertação, também é um tópico que faz parte da computação afetiva.

¹<https://www.nltk.org>

A definição de uma pessoa passa por se estudar a sua personalidade, os seus motivos e seus interesses pessoais (Carneiro 2018, p. 29). A personalidade é um conjunto de características que tornam um indivíduo único. É moldada por várias variáveis, tais como, o percurso feito na vida de uma pessoa e o ambiente em que a pessoa está envolvida (Allport 1961). O estudo da personalidade é retratado como o estudo de atitudes e comportamentos.

Para que houvesse uma maneira de determinar ou definir uma personalidade, foram desenvolvidos vários modelos. Ao longo da pesquisa efetuada durante a dissertação, foram encontrados dois modelos principais de personalidade. Modelos esses, que serão expostos ao longo desta subsecção.

The big five

Segundo (Oliver P. e Sanjay 1999), o modelo de personalidade *The Big Five* (OCEAN (*Openness, Conscientiousness, Extraversion, Agreeableness, Neuroticism*) ou ainda chamado *The Five Factor Model*) é o modelo, sobre a definição de personalidade, que tem mais estudiosos.

Tabela 2.2: Modelo *The big five* adaptado de P. J. Howard e J. M. Howard (1995)

Nível	Baixo	Médio	Alto
Neuroticismo N1: Ansiedade N2: Hostilidade N3: Depressão N4: Autoconsciência N5: Impulsividade N6: Vulnerabilidade	Resistente(N-) Calmo, corajoso Amigável, não se ofende Com esperança, otimista Seguro, à vontade Resiste aos desejos Calmo, resistente	Recetivo (N) Preocupado, calmo Alguma irritabilidade Ocasionalmente triste Por vezes fica embaraçado Às vezes cede aos desejos Algum stress	Reativo (N+) Tenso, medroso Irritável, frustrável e zangado Sem esperança, triste Envergonhado, embaraçado Incapaz de resistir aos desejos Nervoso
Extroversão E1: Acolhimento E2: Gregariedade E3: Assertividade E4: Atividade E5: Procura de excitação E6: Emoções positivas	Introvertido(E-) Frio, normal Evita multidões, solitário Evita afirmar-se Sem pressa Cauteloso Plácido, sério	Ambivertido (E) Atencioso, amável Sozinho ou com os outros Tenta afirmar-se Passo médio Ocasionalmente precisa de emoções fortes Por vezes é alegre e divertido	Extrovertido (E+) Amigável, conservador, afetuoso Gregário, alegre, social Dominante Energético Exibicionista, aprecia estímulos Alegre, espirituoso, divertido
Abertura O1: Fantasia O2: Estética O3: emoções O4: Ações O5: Ideias O6: Valores	Defensor(O-) Realista, prático Insensível à beleza Leque limitado de emoções Prefere o familiar Pragmático Dogmático	Recetivo (O) Pouco imaginativo Aprecia a estética Recetivo a emoções Uma mistura Alguma curiosidade Moderado	Reativo (O+) Imaginativo Valoriza a estética Emotivo, sensível, empático Procura novidade Curioso Tolerante
Amabilidade A1: Confiança A2: Retidão A3: Altruísmo A4: Conformidade A5: Modéstia A6: Sensibilidade	Desafiador(A-) Cínico Calculista Centrado em si Contestador Arrogante Realista, racional	Negociador (A) Prudente Diplomático Disposto a ajudar Moderado Moderado Por vezes influenciado por emoções	Adaptável (A+) Confiável Franco e frontal Gosta de ajudar Complacente Humilde Guiado por emoções ao ajuizar
Conscienciabilidade C1: Competência C2: Ordem C3: Obediência, dever C4: Luta, realização C5: Autodisciplina C6: Deliberação	Flexível(C-) Sente-se incapaz Desleixado Irresponsável Não ambicioso Frustrado, desiste em face da frustração Espontâneo	Equilibrado (C) Preparado Meio organizado Tem em conta as prioridades Procura ter êxito Misto de trabalho e diversão Pensativo	Concentrado (C+) Sente que é capaz Organizado Adesão a padrões de conduta Atraído pelo êxito Percistente e focado no trabalho Cauteloso

Segundo P. J. Howard e J. M. Howard (1995), o *Five Factor Model* (FFM) é composto por cinco características, que representam a personalidade de uma pessoa. Essas cinco características são:

- *Openness* - Que significa a abertura que um indivíduo tem à experiência, ou seja, se gosta de viver novas aventuras. Nesse sentido, uma pessoa com pouca abertura é mais conservadora.
- *Conscientiousness* - Esta característica é relativa à autodisciplina que uma pessoa possui, quando se depara com um obstáculo ou uma dificuldade.
- *Extraversion* - A extroversão determina a capacidade de socializar, comunicar, entre outros.
- *Agreeableness* - A amabilidade de um indivíduo está relacionada com a compaixão, a concordância, a cooperação e a atenção do mesmo.
- *Neuroticism* - O neuroticismo reflete a instabilidade emocional de um indivíduo. As características mais dominantes são a depressão, a negatividade, a raiva e o nervosismo.

A obra P. J. Howard e J. M. Howard (1995) argumenta que na psicologia contemporânea, a personalidade é especificada como uma função de trinta atributos onde cada um é denominado como faceta de personalidade. Na tabela 2.2 são apresentadas as seis facetas dos cinco grupos de características de personalidade (*Openness*, *Conscientiousness*, *Extraversion*, *Agreeableness*, *Neuroticism*). Onde o O é a abertura à experiência, o C é a consciencialização, o E é a extroversão, o A é a amabilidade e o N é o neuroticismo.

Modelo PEN

O modelo PEN é uma adaptação do modelo FFM. Mas com apenas três grupos, em vez de cinco. Eysenck (1991) afirmou que o modelo FFM tinha grupos que significam o mesmo.

Portanto, os três grupos do modelo PEN são: psicoticismo, extroversão e neuroticismo. Estes três grupos estão ao mesmo nível. O psicoticismo corresponde a características como: episódios psicóticos e comportamentos agressivos. Os grupos da extroversão e do neuroticismo correspondem aos mesmos grupos do modelo FFM.

2.4.2 Análise de emoções

O conceito de emoção diferencia-se do conceito de sentimento. O sentimento é um resultado de uma experiência passada, sendo um estado neuropsicológico. Uma emoção é um impulso nervoso que despoleta uma ação (Freitas et al. 2013). Existem várias teorias de como se formam as emoções: a psicológica; a neuronal e a cognitiva. A teoria cognitiva determina que as interações do ser humano com o ambiente onde se encontra inserido, provocam emoções. Esta teoria será a mais valorizada no seguimento desta dissertação.

Como já referido, as emoções são relevantes para os processos de tomada de decisão, bem como para auxiliar um computador a obter os melhores resultados.

Como Gross e Thompson (2007) refere no seu trabalho, uma emoção é uma sensação que nos impulsiona a fazer algo. Tal como foram desenvolvidos modelos para determinar e

classificar a personalidade dos indivíduos, também se desenvolveu modelos para determinar e classificar emoções.

Esses modelos são uma grande ajuda para determinar emoções nas ciências de computação.

O modelo OCC

O modelo OCC, desenvolvido por *Ortony, Clore and Collins* (OCC), é o modelo mais utilizado em ciências de computação para descrever o processo da criação de emoções. O modelo dá, de forma clara, uma estrutura das condições e variáveis que afetam as intensidades das emoções (Steunebrink et al. 2009).

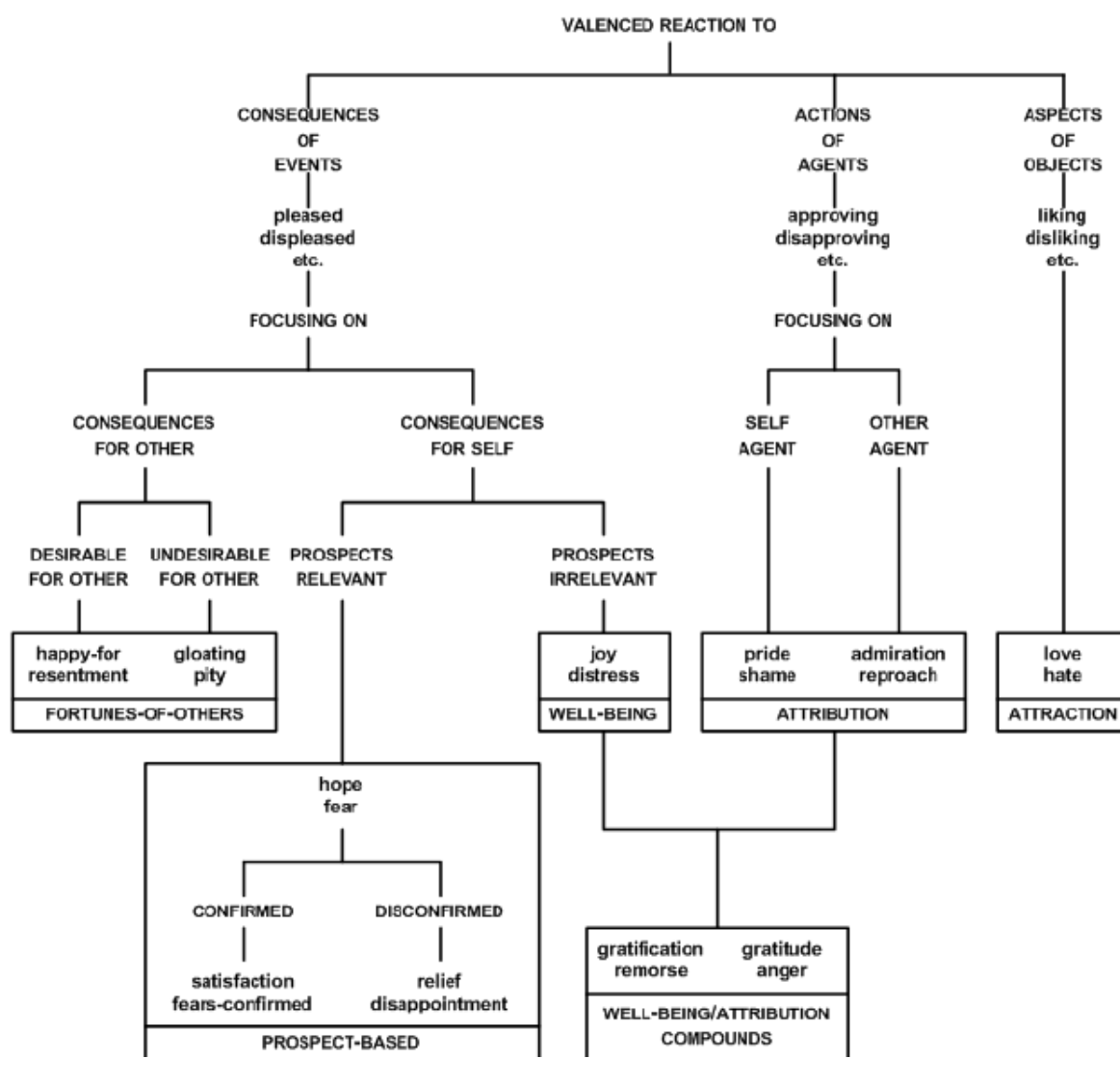


Figura 2.5: Modelo OCC adaptado de Ortony, Clore e Collins (1990)

Segundo o modelo em estudo, uma emoção é gerada como uma reação a uma consequência de eventos, ações ou aspectos de um objeto. Para calcular a intensidade de uma emoção são consideradas variáveis globais: o senso de realidade e proximidade, as variáveis locais,

a probabilidade de algo acontecer, o esforço para atingir um objetivo e a possibilidade de atingir o objetivo. As percepções dessas emoções estão relacionadas com objetivos, padrões e preferências de um indivíduo. O modelo OCC é ilustrado na figura 2.5.

O modelo OCC é bastante utilizado em ciências da computação, para analisar emoções. Tal como foi apurado na pesquisa, muitos dos investigadores basearam os algoritmos de detecção e gerenciamento de emoções no modelo OCC (Carneiro 2018; Loureiro, Marreiros e Neves 2011).

Modelo emocional de Ira Roseman

Segundo Roseman, Spindel e Jose (1990), uma emoção é gerada através de um procedimento associação de eventos. Os eventos são classificados de várias maneiras:

- Podem ser por motivos consistentes ou sem motivos consistentes. Os eventos por motivos consistentes são os eventos que vão de encontro com os objetivos da pessoa. Os eventos sem motivos consistentes não têm em consideração os objetivos da pessoa;
- Podem ser classificados pela causa. Se foram causados individualmente, por circunstâncias ou por outra pessoa;
- Um evento também pode ser iniciado de forma a fugir de uma consequência ou devido a uma recompensa;
- Por fim, um evento pode ser classificado se for um evento certo, um evento que não se estava à espera ou se aconteceu de maneira incerta.

Hourglass of emotions

Um sentimento é criado através de uma ou várias emoções. Um sentimento depende não só das emoções presentes, mas também da intensidade das mesmas.

Inspirado no trabalho de Plutchik (2001), sobre as emoções humanas, o *hourglass of emotions* é um modelo de classificação de emoções (Cambria, Livingstone et al. 2012). O modelo organiza as seguintes quatro emoções primárias independentes:

- Sensibilidade - Se um indivíduo está confortável com as dinâmicas de interação;
- Atenção - Se um indivíduo está interessado nas interações;
- Aptidão - Se um indivíduo é confiante nos benefícios das interações;
- Agradabilidade - Se um indivíduo está contente com interações.

Cada uma destas quatro emoções primárias (ou quatro dimensões de afeição) se dividem em seis emoções: três positivas e três negativas, cada uma delas com diferentes intensidades. Como se pode verificar na figura 2.6.

A sensibilidade divide-se em três emoções positivas (ira, raiva e aborrecimento) e em três emoções negativas (apreensão, medo e terror).

A atenção divide-se em emoções positivas (vigilância, antecipação, interesse) e emoções negativas (distração, surpresa e espanto).

Quanto à aptidão é dividida em emoções positivas (admiração, confiança e aceitação) e emoções negativas (tédio, nojo e repugnância).

Por fim, a agradabilidade divide-se em três emoções positivas (êxtase, alegria e serenidade) e em três emoções negativas (pensativo, tristeza e luto).

O modelo apresentado é bastante utilizado em sistemas de análise de sentimento, como nos trabalhos (Cambria, Olsher e Rajagopal 2014; Cambria, Poria et al. 2018). A função que determina as emoções é composta pela soma do grau de cada uma das quatro emoções primárias: Sensibilidade, representada por S ; Atenção, representada por At ; Aptidão, representada por Ap e Agradabilidade, que é representada por P . O grau varia de -1 a 1. O resultado da soma é dividido por três, vezes o número total de conceitos. A função que determina a polaridade é a que está presente na equação 2.1.

$$p = \sum_{i=1}^N \frac{P(ci) + |At(ci)| - |S(ci)| + Ap(ci)}{3N} \quad (2.1)$$

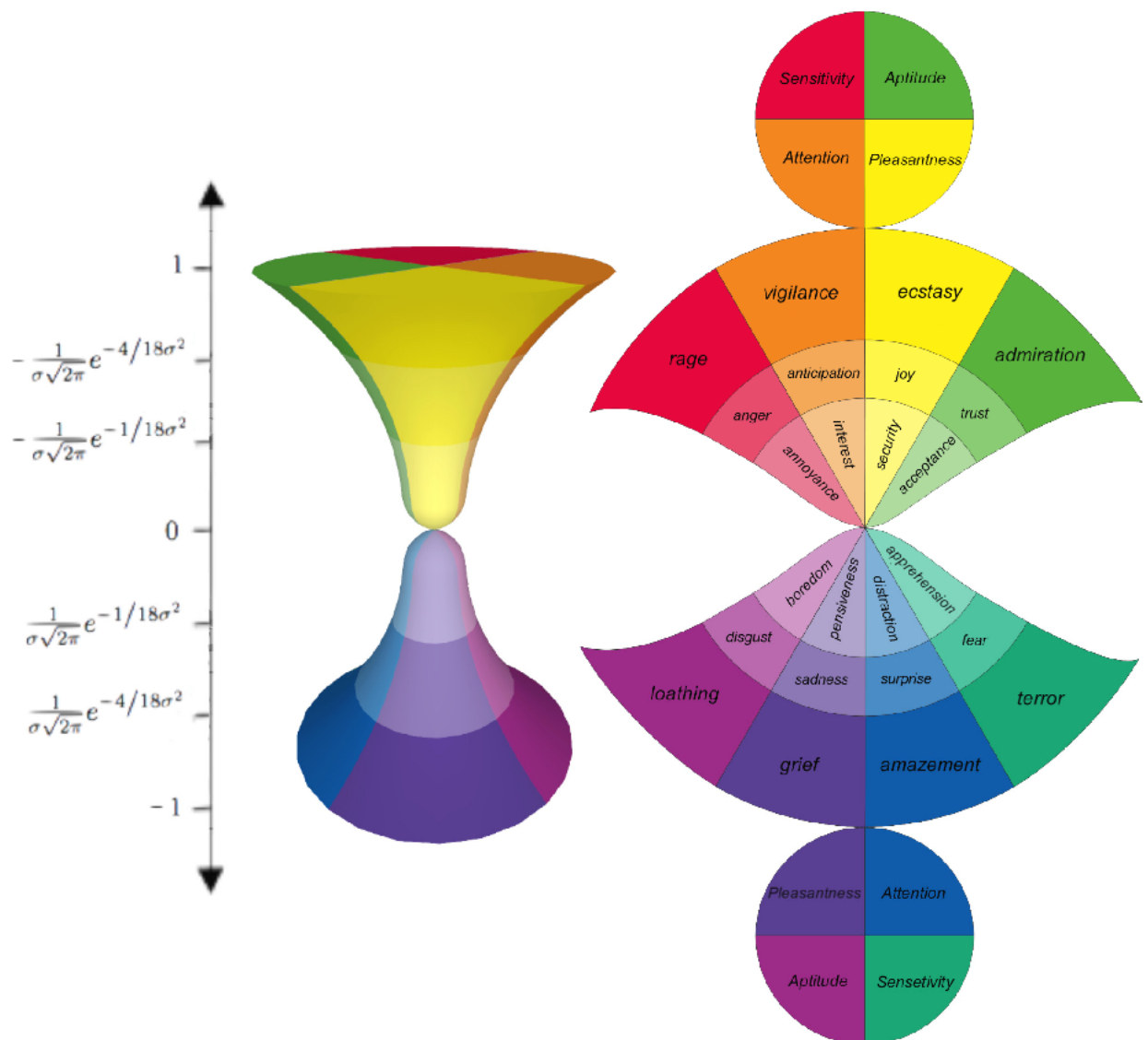


Figura 2.6: Modelo *Hourglass of emotions* adaptado de Cambria, Livingstone et al. (2012)

2.4.3 Datasets e lexicons

Os algoritmos necessitam de informação para funcionarem. Essa informação, dependendo da função dos algoritmos, irá variar. No caso de algoritmos de processamento de linguagem natural ou utilizados na computação afetiva, necessitam de palavras ou expressões para poderem ser testados e avaliados.

Uma *dataset* é um conjunto de dados organizados sobre algo, que deve preferencialmente ser do tema.

Um *lexicon* é uma base de dados que é uma espécie de dicionário. Para o problema em questão, o *lexicon*, deve ter as palavras que são significativas para a análise de emoções, bem como as respectivas cotações de cada umas delas.

Para avaliar os algoritmos a serem desenvolvidos há necessidade de se utilizar *dataset* e *lexicon*. Para conhecer os *datasets* e *lexicons* foi efetuada uma pesquisa. Nessa pesquisa foram encontrados os *datasets*: *Imdb*; *tweet*; *UN* e *US Debate* e os *lexicons*: *WordNet* e *SenticNet*.

Imdb dataset

A *imdb*, que significa *internet movies database*, é uma base de dados com informações sobre filmes e séries televisivas. Informações sobre o elenco, um resumo, o *trailer*, entre outros, contendo, também, comentários sobre esse filme ou série.

A *Imdb dataset* contém cerca de 50000 revisões, tendo cada filme um máximo de 30 revisões. Além disso, estão divididas em 25000 revisões positivas e 25000 revisões negativas. Uma revisão é considerada negativa se lhe for atribuída quatro ou menos estrelas e positiva de 7 estrelas ou superior a 7. As restantes são consideradas neutras.

O contexto deste *dataset* não é o mesmo do problema, que é a decisão em grupo. Porém, foi utilizado por vários autores para testarem algoritmos de análise sentimental, como por exemplo (Hogenboom et al. 2014; Maas et al. 2011).

Tweeter dataset

O *tweeter* é uma rede social, na qual as pessoas, entre muitas outras coisas, podem fazer comentários e outro tipo de publicações.

A *Tweeter dataset* é um conjunto de dados organizados por autor, publicação ou *tweet* e data da publicação.

O contexto desta *dataset* não é o mesmo do problema, que é a decisão em grupo. Porém, foi utilizado por vários autores para testarem algoritmos de análise sentimental, como por exemplo, (Agarwal et al. 2011; Barbosa e Feng 2010; Patodkar e I.R 2016).

UN dataset

As *United nations* ou UN é um grupo de estados, cujos representantes se reúnem, uma vez por ano, desde 1947, para discutir diversos tópicos.

A *UN dataset* contém todas as afirmações, de cada representante, na assembleia da UN, desde 1970 até 2015. A informação está organizada por país, ano, número da sessão e o próprio discurso.

O contexto da informação desta *dataset* tem algumas semelhanças com a decisão em grupo, visto que são argumentações entre pessoas. O único trabalho encontrado na pesquisa que utilizou esta *dataset* era sobre política, não da área de ciência de computação.

US debate dataset

Em 2016, como é habitual em países democráticos, nos Estados Unidos da América houve debates entre os concorrentes à presidência americana.

A *US debate dataset* contém todos os debates, entre todos os deputados do partido republicano e democrático. A informação está dividida em argumentos e em quem o disse.

WordNet

WordNet é uma base de dados léxica da língua inglesa. Como é afirmado no *WebSite*²: "Que contém nomes, adjetivos e verbos agrupados por conjuntos de sinónimos". Para além disso, também são agrupadas por sentido.

O *lexicon* já contém uma API, em que, apresentado uma expressão, devolve o significado dessa palavra, ou seja, o grupo de sinónimos a que pertence.

Este *lexicon* inspirou à criação de outros, como o *SenticNet*, que será referido no subcapítulo seguinte, bem como no desenvolvimento de trabalhos na área de computação afetiva.

SenticNet

SenticNet, como referido no website oficial,³ "pode ser considerada uma base de conhecimento ou uma *framework*". A pesquisa foca-se mais na base de conhecimento.

A base de conhecimento da *SenticNet* tem cada expressão ou palavra agrupada por conceito. O *output*, que esta base de conhecimentos fornece, é composto pela polaridade que é dada por um valor entre -1 a 1, bem como pela emoção presente no *input*. Esta emoção está avaliada em quatro dimensões (Agradabilidade, Atenção, Sensibilidade e Aptidão). É uma ferramenta grátis e pode ser utilizada como API ou em ficheiro *xml*.

Até esta data, o *lexicon SenticNet* foi utilizado por diversos autores em diversos trabalhos sobre análise de emoções, tais como (Agarwal et al. 2011; Barbosa e Feng 2010; Loureiro, Marreiros e Neves 2011).

SenticNet também tem, à disposição dos utilizadores, uma *framework*. Esta *framework* possui módulos, que lidam com o tratamento de texto e classificação das emoções. Todas as emoções estão de acordo com o modelo *Hourglass of emotions*.

2.4.4 Abordagens utilizadas para determinar e classificar emoções

A análise de emoções em texto é efetuada por vários algoritmos, nos quais estão presentes diferentes tipos de abordagens. Nesta secção iremos falar do estado de arte, relativo às abordagens utilizadas por diversos autores, encontradas durante a fase de pesquisa. Essas abordagens utilizam *Natural language processing* (NLP) de modo a analisarem um texto subjetivo.

O termo NLP designa um conjunto de técnicas que conseguem analisar o texto de linguagens dos seres humanos. Estas técnicas são utilizadas, cada vez mais, em inteligência artificial ou em *machine learning* (Bird, Klein e Loper 2009).

²<https://wordnet.princeton.edu/>

³<https://sentic.net/>

Nesta subsecção vão ser expostos os diferentes algoritmos que servem para classificar e detetar emoções: agentes inteligentes, o classificador linear SVM, um classificador *Naive Bayes* e a *framework SentiNet*.

Agentes inteligentes

Agentes inteligentes são entidades autónomas que conseguem perceber o ambiente através de sensores e interpretar informações recebidas, executando uma ação de acordo com essa interpretação. Os dois estudos seguintes utilizam esta metodologia para obter emoções.

O estudo feito por Marreiros, Santos et al. (2010) apresentou um SADG que utiliza agentes inteligentes para auxiliar indivíduos de uma determinada reunião. O modelo OCC foi utilizado na determinação das emoções. A arquitetura está dividida em três camadas: conhecimento; razão e interação. Na camada do conhecimento, o agente tem os perfis de todos os participantes da reunião, as suas preferências e os seus objetivos. Na camada da razão existe um módulo de argumentação, que cria argumentos. O módulo existente na camada de decisão, ajuda na tomada de decisão, selecionando o melhor argumento. O módulo de emoções, gera emoções e o humor do participante. Por fim, a camada de interação serve como interface para o simulador do SADG e trata da comunicação entre os diversos agentes. As diferentes emoções têm diferentes intensidades e tempos diferentes no seu desaparecimento, ou seja, cada tipo de emoção leva um determinado tempo até se desvanecer por completo.

No artigo Santos et al. (2011) estudou-se o contexto dos SADG, utilizando uma arquitetura multiagentes. A grande novidade deste trabalho foi dar a possibilidade de moldar a personalidade de um agente. A personalidade é moldada conforme o modelo FFM. Para tal, cada agente tem que responder ao questionário do *Big Five Inventory*. Esse questionário contém 44 questões e cada uma será respondida na forma de uma escala de *Likert*. Os agentes poderão ou não saber a personalidade um do outro. O foco deste trabalho é fazer agentes que simulem uma argumentação. As emoções de cada agente são importantes para a decisão. Concluiu-se que os agentes chegavam mais depressa a acordo quando conheciam as personalidades uns dos outros.

SVM

O *Support Vector Machine* ou SVM é um algoritmo que pode servir como classificador linear. Este divide diferentes classes de informação, de forma a diferenciar os tipos de informação. Quando aplicados em algoritmos, consegue-se classificar informação de acordo com algum tipo de critério (Aggarwal e Zhai 2012). Para que seja possível a classificação da informação utiliza-se hiperplanos. Cada hiperplano equivale a uma classificação. De modo a criar esses hiperplanos a informação é transformada em pontos. Após essa transformação, os pontos funções quadráticas são designadas para criar esses hiperplanos. Deve-se verificar, sempre, se os pontos que representam a informação nova estão no mesmo hiperplano da informação já classificada. Se os pontos que representam a informação nova estiverem num hiperplano já conhecido irão pertencer à classificação que esse hiperplano representa.

CRF

O *Conditional Random Fields* (CRF) é um modelo probabilístico, utilizado para classificar. "A conditional random field (CRF) is a special case of Markov random field, but each state of node is conditional on some observed values".

(Aggarwal e Zhai 2012, p. 261). Este modelo usado em algoritmos pode classificar informação.

Antes de se poder explicar este algoritmo é necessário explicar o que são modelos gráficos, como o modelo de *Markov Random Fields*. O modelo de *Markov Random Fields* é um conjunto de grafos que une variáveis aleatórias. Esses grafos decidem a dependência entre as diversas variáveis. E cada variável está representada por um nó e cada aresta significa as dependências entre as variáveis.

O modelo de CRF recorre às representações gráficas e às técnicas de classificação discriminatória. Inicialmente o algoritmo agrupa a informação conhecida num grafo. Posteriormente, o algoritmo utiliza métodos para calcular a probabilidade da nova informação, já transformada em nós. Por fim, coloca o nó da nova informação ligada aos nós em que a probabilidade é maior (Sutton 2012).

Naive Bayes classifier

O *naive Bayes classifier* é mais um classificador de informação com base num modelo probabilístico. O classificador tem também como base o teorema *Bayers*. O algoritmo tem uma abordagem, "*naive*", em que todas as características são independentes entre si. Ou seja, a presença de certas características não tem dependência ou relação, como por exemplo, relações entre palavras (Aggarwal e Zhai 2012). O *naive Bayes classifier* não faz comparação entre frases, mas sim entre cada palavra, utilizando uma abordagem *bag of words*. Esta abordagem revela-se mais útil, quando se utiliza pouca quantidade de informação. O teorema aplicado em análise de emoções está apresentado em 2.2, conforme expresso no artigo (Patodkar e I.R 2016).

$$P(s|M) = \frac{P(s) * P(M|s)}{P(M)} \quad (2.2)$$

A equação dita que a probabilidade de um sentimento (s) estar presente numa frase (M) é igual à multiplicação da probabilidade anterior de um sentimento (P(s)) com a probabilidade de uma frase ter um sentimento (P(M|s)). Tudo isto a dividir pela probabilidade de aparecer uma frase (P(M)).

Este modelo tem várias variações. As seguintes são três, das mais utilizadas: (Bird, Klein e Loper 2009):

- *Gaussian* - Quando se assume que as características assumem distribuição normal para a classificação;
- *Multinomial* - Utiliza uma contagem discreta. No caso da NLP, vai contar todas as vezes que uma palavra ocorre num texto, em vez de contar uma palavra de cada vez;
- *Bernoulli* - É um modelo binomial, utilizado quando os vetores de características são binários.

Implementações

No artigo Pang e Lee (2004), os investigadores optaram, para poupar tempo de execução do módulo, por classificar frases pela sua subjetividade. As frases objetivas, ou não subjetivas, não eram consideradas pois não tinham emoções associadas. Após a classificação das frases, os investigadores aplicaram um algoritmo do tipo SVM e um algoritmo do tipo *Naive Bayes classifier*, dos quais se obteve uma polaridade do texto. O *Naive Bayes classifier* foi o mais eficaz, com precisão de 86.4% com cortes mínimos contra os 85.45% sem cortes. Já o SVM teve uma precisão de 86.15% com os cortes mínimos.

No estudo de Ohana e Tierney (2009), os investigadores utilizaram o *lexicon SentiWordNet* para detetar as polaridades das expressões e das palavras. Desenvolveram um algoritmo para verificar se havia negações, utilizando listas com expressões relativas a negação. Estas listas dividem-se em expressões que alteram a polaridade dos termos seguintes, dos termos anteriores ou não alteram a polaridade. Os algoritmos foram todos feitos com a implementação SVM, quer o de verificação de negação, quer o de determinar a polaridade. Conclui-se que a eficácia dos algoritmos desenvolvidos neste trabalho, usando o *lexicon SentiWordNet* como funcionalidade, eram melhores do que os já existentes, como visível na tabela 2.4. Não obstante, a diferença entre se basear num *lexicon* é ligeiramente melhor do que não se basear num *lexicon*.

Tabela 2.3: Comparação entre abordagens adaptado de Pang e Lee (2004)

Model	Accuracy
SentiWordNet – Term Counting (this research)	65.85%
SentiWordNet Scores used as Features (this research)	69.35%
Term Counting - Manually built list of Positive/- Negative words	69.00%
Term counting from Combined Lexicon and valence shifters	67.80%

No artigo Lin et al. (2006), os investigadores utilizaram artigos de um *website* que continha opiniões sobre o confronto da Palestina e o Israel. Utilizaram um algoritmo que analisou a subjetividade das frases, de modo a selecionar as subjetivas. Depois utilizaram-se algoritmos do *Naive Bayes classifier*, do *Bayesian* (ou *Maximum a posteriori*) e algoritmos SVM, para classificar o texto. Como se pode comprovar na tabela.

Tabela 2.4: Comparação entre abordagens adaptado de Lin et al. (2006)

Model	Data Set	Accuracy	Reduction
Baseline		0.5	
SVM	Editors	0.9724	
NB-M	Editors	0.9895	61%
NB-B	Editors	0.9909	67%
SVM	Guests	0.8621	
NB-M	Guests	0.8789	12%
NB-B	Guests	0.8859	17%

Os autores do artigo Agarwal et al. (2011) fizeram análise de emoções a vários *Tweets*. Utilizaram um algoritmo SVM para classificar as emoções encontradas num texto. O algoritmo referido fez duas classificações: positivos e negativos ou positivos, negativos e neutros. Nestas duas classificações, foram comparados diferentes modelos de dados: um monograma, um modelo de árvore *Kernel* e um modelo baseado em 100 *senti-features*.

Um monograma é um *n-gram*, porém só possui um conjunto de repetições de uma sequência. Este modelo é dos mais utilizados em análise de emoções.

O modelo baseado em 100 *senti-features* é um modelo que divide o texto em 100 características para análise.

E o modelo de árvore *Kernel* é um modelo não linear, que procura parecenças com características, alinhando-as em forma de árvore.

Os mesmos algoritmos SVM foram utilizados para realizar uma classificação para cada tipo de modelo e usaram-se testes de *5-fold cross-validation* para a sua validação. Nas classificações provou-se que o modelo de árvore *Kernel* foi o mais eficiente. O *lexicon* utilizado foi *WordNet*, como se pode verificar nos resultados demonstrados nas tabelas, 2.5 e 2.6 .

Tabela 2.5: Resultados de teste *2-cross evaluation* adaptado de Agarwal et al. (2011)

Modelo	Avg. Acc (%)	Std. Deviation
Unigram	71.35	1.95
Senti-features	71.27	0.65
Kernel	73.93	1.50
Unigram + Senti-features	75.39	1.29
Kernel + Senti-features	74.61	1.43

Tabela 2.6: Resultados de teste *3 way evaluation* adaptado de Agarwal et al. (2011)

Modelo	Avg. Acc (%)	Std. Deviation
Unigram	56.58	1.52
Senti-features	56.31	0.69
Kernel	60.60	1.00
Unigram + Senti-features	60.50	2.27
Kernel + Senti-features	60.83	1.09

No artigo Cambria, Olsher e Rajagopal (2014), os autores decidiram aplicar um estudo sobre a ferramenta *SenticNet* na versão 3.0. *SenticNet* não é apenas um *lexicon*, mas também uma ferramenta que trata frases, pode dar a sua polaridade, entre outras informações. Para que testassem a ferramenta, os autores utilizaram históricos de pacientes do sistema nacional de saúde britânico. O *SenticNet* utiliza uma abordagem de conceitos, pois em vez de classificar por palavras, classifica por conjuntos de palavras com um sistema de grafos. Assim, é permitido que se tenha em conta a negação e a identificação de conceitos do senso comum. Conceitos, estes, que não são encontrados em algoritmos que utilizam a *bag of words*, abordagem com menos eficácia. Na tabela 2.7 pode-se encontrar a comparação da precisão entre as três versões da ferramenta *SenticNet*, em diferentes áreas.

Tabela 2.7: Comparação entre as precisões diferentes versões de adaptado de Cambria, Olsher e Rajagopal (2014)

Category	SenticNet	SenticNet 2	SenticNet 3
clinical service	59.12%	69.52%	78.06%
communication	66.81%	76.35%	80.12%
food	67.95%	83.61%	85.94%
parking	63.02%	75.09%	79.42%
staff	58.37%	67.90%	76.19%
timeliness	57.98%	66.00%	75.98%

No artigo Patodkar e I.R (2016), os investigadores desenvolveram um trabalho sobre a análise de emoções a *tweets*. Foi utilizada uma *Application Programming Interface* (API) do *tweeter* para construir uma dataset com o valor emocional de cada um dos *emoticons*. Após a análise de texto, foram construídos *n-grams*. Os *n-grams* são repetições de uma ou várias sequências de um determinado texto. Pode representar palavras ou sílabas. Depois de se formar os *n-grams*, procedeu-se à classificação das emoções presentes. Para além de utilizarem os *n-grams*, também utilizaram os *POS-tags*, que são apontadores. Para a classificação foram usados diferentes tipos de algoritmos: *Support Vector Machine* (*Support vector machines* (SVM)), *Conditional Random Fields* (*Conditional Random Fields* (CRF)), *naive Bayes classifier*. O *naive Bayes classifier* é o mais eficaz. Não foi especificado o modelo de emoções utilizado. Os resultados dos testes executados pelos autores determinaram os valores de precisão dos diferentes algoritmos, como é exposto da tabela adaptada do trabalho em análise.

O Cambria, Poria et al. (2018) é um artigo que explica as alterações da nova versão da *framework* e *lexicon SenticNet*, versão 5.0. Nesta nova versão é utilizado um modelo biLSTM para transformar as frases recebidas. Depois, a informação na forma de LSTM vai ser analisada e agrupada em conceitos simples, como é visível na figura 2.7. Na versão do SenticNet são verificadas possíveis contradições. Esta *framework* contém um módulo chamado *AffectiveSpace*, responsável na determinação da polaridade de cada conceito. A *AffectiveSpace* utiliza um *lexicon* de senso comum e a abordagem de *Johnson and Lindenstrauss's Lemma* para reduzir o tamanho do vetor com o senso comum. Desta forma, melhora a *performance* do algoritmo. Essa abordagem está presente na fórmula 2.3. Nesta fórmula, a variável x é um conjunto de vetores num espaço Euclidean; d é a dimensão original do espaço Euclidean; m é a dimensão do espaço, para o qual é pretendido reduzir a informação; o ϵ é o parâmetro da tolerância que representa o máximo de permitido de distorção; e o ϕ é uma matriz aleatória.

$$\sqrt{\frac{m}{d}} \|x - y\|_2 (1 - \epsilon) \|\phi x - \phi y\|_2 \leq \sqrt{\frac{m}{d}} \|x - y\|_2 (1 - \epsilon) \quad (2.3)$$

Para a avaliação desta *framework* foram utilizados dois *datasets*: o *Blitzer* e um sobre revisões de filmes. A precisão foi 94.6%.

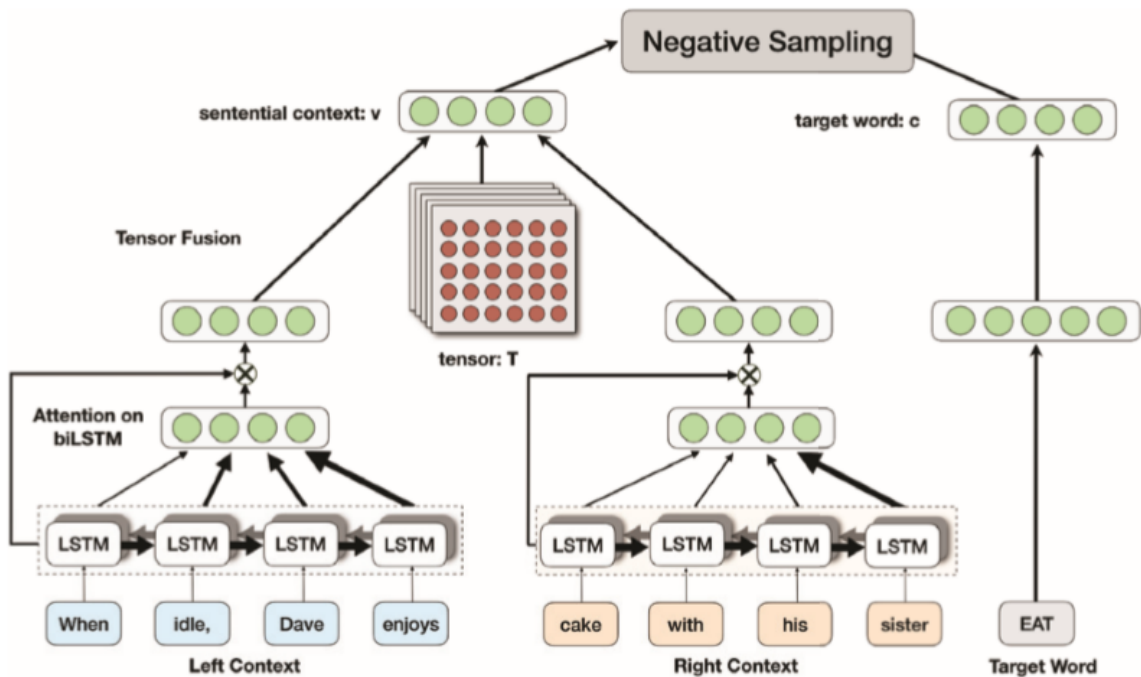


Figura 2.7: Modelo de LST adaptado de Cambria, Poria et al. (2018)

2.4.5 Impacto emocional em terceiros

Nesta subsecção iremos expor a pesquisa efetuada sobre o impacto emocional. Um dos objetivos deste estudo é determinar o impacto que as emoções podem ter nas restantes pessoas que intervêm na tomada de decisão.

Ao argumentar, um indivíduo pode transmitir emoções ao recetor, mesmo de forma inconsciente (Bersade 2002; Marreiros, Ramos e Neves 2005; Martinovski e Mao 2009; Tsai et al. 2020). Este processo é denominado de contágio emocional. Este contágio emotivo tem impacto na tomada de decisões.

Estudo de Bersade

No estudo de Bersade (2002) foi apresentado um modelo de contágio emocional. Este investigador demonstra que as emoções negativas têm mais facilidade em serem absorvidas por terceiros do que as emoções positivas. Não obstante, as emoções positivas proporcionam a uma maior cooperação. Através de experiências entre duas pessoas, o investigador comprovou que o impacto emocional varia conforme o grau de amenidade ou prazer e conforme a energia das emoções. Dividiu os indivíduos em quatro condições experimentais:

- *Cheerful Enthusiasm* - Quando o indivíduo é muito expressivo, exprime-se com muita energia e expressa emoções positivas;
- *Hostile Irritability* - Quando o indivíduo é muito expressivo, exprime-se com muita energia, e expressa emoções negativas;
- *Serene Warmth* - Quando a pessoa se expressa com pouca energia e expressa emoções negativas;
- *Depressed Sluggishness* - Quando a pessoa se expressa com pouca energia e expressa emoções negativas.

Em suma, as emoções como a felicidade, a surpresa, a admiração e o interesse ajudam a que haja concordância e cooperação. As emoções como a raiva, o medo, o desgosto e a tristeza dificultam cooperação.

Modelo EASI

Numa argumentação, os intervenientes podem ou não estar inseridos num ambiente cooperativo ou competitivo. Segundo G. van Kleef, C. De Dreu e A. Manstead (2010), o impacto que uma emoção pode ter num determinado indivíduo depende, também, se esse indivíduo se encontra num ambiente competitivo ou cooperativo. Neste estudo, o modelo *Emotions as Social Information* (ou EASI) foi desenvolvido. Dentro deste modelo, existe a informação de que as emoções podem ajudar a aumentar ou diminuir o grau de cooperação. O impacto emocional no processo da tomada de decisão vai variar em função do ambiente competitivo ou cooperativo, em que os participantes se encontram, e da emoção que eles experienciam. Se forem utilizadas emoções positivas, num ambiente cooperativo, é provável que haja um aumento na cooperação. Mas, se forem utilizadas emoções positivas, num ambiente competitivo, a competição poderá aumentar. Caso estejam presentes emoções de frustração, como a raiva, num ambiente cooperativo, a cooperação poderá sofrer uma

redução. A frustração, em ambientes competitivos, pode levar a um aumento da cooperação. A tristeza, num ambiente cooperativo, pode aumentar a cooperação e num ambiente competitivo aumentar a competição. Num cenário de culpa, num ambiente cooperativo, a competição diminui. Mas, num ambiente competitivo, o cenário de culpa provoca uma maior competição. Como é visível na tabela 2.8.

Tabela 2.8: Comportamentos de diversas emoções em ambientes cooperativos e competitivos G. van Kleef, C. De Dreu e A. Manstead (2010)

Emoção	Sinal Social	Ambiente cooperativo	Ambiente competitivo
Felicidade, alegria, contentamento	Oportunidade, afiliação	Aumento da cooperação	Aumento da competição
Raiva, frustração, irritação	Dominância, agressividade	Redução da cooperação	Aumento da cooperação
Tristeza, preocupação	Suplicação	Aumento da cooperação	Aumento da competição
Culpa	Apaziguamento	Redução da competição	Aumento da competição

Estudo de impacto de emoções em negociações de Van Kleef

Nos artigos (Van Kleef, De Dreu e Manstead 2006; G. A. van Kleef, C. K. W. De Dreu e A. S. R. Manstead 2004) as emoções são relacionadas com a escolha da estratégia ou com a reação de intervenientes em negociações. No artigo de G. A. van Kleef, C. K. W. De Dreu e A. S. R. Manstead (2004), os investigadores estudaram o efeito da raiva contra o efeito da felicidade em todos os intervenientes de uma decisão. Estes estudos concluíram que o impacto emocional não depende apenas da raiva e de felicidade, mas também se os indivíduos se encontram em ambientes competitivos (baixa confiança) ou em ambientes cooperativos (alta confiança). Tanto num ambiente cooperativo, como competitivo, se um indivíduo transmitir felicidade, existe uma maior probabilidade dos restantes intervenientes fazerem exigências ou cedências. Caso, um indivíduo expresse raiva, num ambiente cooperativo, os restantes intervenientes diminuam as exigências e há uma menor probabilidade de cedência. Já, num ambiente competitivo, os intervenientes aumentam as exigências e diminuem o nível de cedências.

No artigo G. van Kleef, C. De Dreu e A. Manstead (2010) os impactos da culpa e da tristeza, nas negociações, são estudados. Neste estudo, concluiu-se o mesmo que no anterior: o impacto de uma emoção também depende do ambiente em que todos os intervenientes estão inseridos. Se o indivíduo sentir culpa, num ambiente cooperativo, o grau de exigência aumenta e a concordância torna-se difícil. O contrário se verifica com a tristeza. Num ambiente de baixo nível de confiança (ambiente competitivo), a culpa aumenta o nível de exigência. Já a tristeza, diminui o nível de exigência. O resultado destes dois estudos pode ser observado na tabela 2.9.

Tabela 2.9: Comportamentos de diversas emoções em ambientes cooperativos e competitivos

Emoção	Ambiente Cooperativo	Ambiente competitivo
Felicidade, alegria, contentamento	Aumento da cooperação e aumento de cedência	Aumento do nível de exigência e maior facilidade de cedência
Raiva, frustração, irritação	Redução de exigência e diminuição de cedência	Aumento de exigência e diminuição de cedência
Tristeza, preocupação	Diminuição de nível de exigência	Diminui o nível de exigência
Culpa	Aumenta o nível de exigência	Aumenta o nível de exigência

2.5 Conclusões e discussões

Nesta secção foram analisados conceitos relacionados com os temas: processo de decisão, computação afetiva e impacto de emoções. Na tomada de decisão explicamos a importância das decisões, bem como as vantagens dos processos de tomada de decisão em grupo e da sua ligação com as emoções. Na parte da computação afetiva foram analisados, com a ajuda da psicologia, diversos conceitos e modelos sobre a personalidade de uma pessoa e as suas emoções. Nesta secção ainda descortinamos alguns estudos relativos à área de computação, que utilizaram esses conceitos e modelos.

Detetamos que, ao contrário da análise de emoções, não existem muitos estudos sobre a análise sentimental em texto, no contexto dos sistemas de apoio à tomada de decisão em grupo.

A biblioteca *NLTK* foi uma ajuda na compreensão dos vários textos analisados, neste módulo. Esta biblioteca só é compatível com a linguagem *Python* e é das mais rápidas e eficazes ferramentas para processamento de linguagem natural. Por essa razão, optamos pela sua utilização.

Como verificado na subsecção 2.4.4, falamos dos estudos que vários investigadores desenvolveram para analisarem as emoções e os sentimentos de um texto. Cada um desses estudos apresentam os seus benefícios e as suas falhas. Por exemplo, o algoritmo SVM é eficiente em muitas dimensões, quando estas são maiores que o número de amostras. A framework de análise de emoções, que mais se revelou útil para a realização desta dissertação, foi a *framework SenticNet*. A razão para a sua escolha é óbvia. Trata-se de uma framework com maior precisão na deteção e determinação das emoções presentes num texto.

Capítulo 3

Design e implementação da solução

Neste capítulo iremos apresentar a solução encontrada para o problema apresentado nesta dissertação e iremos explicar o seu *design*.

A solução proposta passa pela criação de um módulo para a determinação de emoções presentes em texto, bem como a determinação do impacto causado por eles.

Na secção 3.1 iremos desenvolver a arquitetura do sistema, bem como a arquitetura do módulo de análise de emoções, da solução proposta.

3.1 Arquitetura do sistema

A arquitetura é a organização de um sistema, dos seus componentes e da relação entre eles. O sistema da solução vai seguir o *Single-Responsability Principle*, princípio da responsabilidade.

O sistema da solução apresenta uma arquitetura de camadas, que consiste na separação do sistema por camadas hierárquicas, em que as de baixo não comunicam com as de cima. Mas, as de cima comunicam com as de baixo. As camadas são as seguintes:

- *Data Access Layer (Data Access Layer (DAL))* – É a responsável pelas operações Operações Create, Remove, Update, Delete (CRUD) com a base de dados, ou seja, responsável por todos os dados do sistema;
- Módulo de análise de emoções – Camada que vai comunicar com a DAL, de modo a obter dados e fazer a análise de emoções.



Figura 3.1: Modelo de componentes do sistema

A DAL é uma camada que não tem lógica de negócio, apenas retorna os dados necessários para a análise de emoções.

A figura 3.1 contém os diversos componentes do sistema : *SenticNetDB*, *SenticNetAPI* e *AnalyseModule*. A base de dados (*SenticNetDB*), tem os conceitos de várias línguas armazenados, bem como as suas polaridades. A API do *SenticNet* (*SentiNetAPI*), é responsável pela comunicação com a base de dados. Quer a *SenticNetDB* quer a *SentiNetAPI*, vão fazer parte da camada inferior, a DAL. Por fim, o módulo de análise de emoções (*AnalyseModule*) irá conter todos os algoritmos e métodos para analisar as emoções.

Por sua vez, o componente do módulo de análise de emoções vai-se subdividir em vários módulos. O primeiro módulo, *TextAnalyseModule*, receberá um texto, analisá-lo-á e enviará as palavras desse mesmo texto para a API do *SenticNet*. Posteriormente, irá receber o resultado da API. Os dados resultantes deste módulo irão para outro módulo, *SentimentModule*, que irá verificar o impacto emocional. O *SentimentModule*, imprimirá os resultados: todas as emoções e o impacto final. Os componentes do módulo de análise de emoções podem ser observados na figura 3.2.

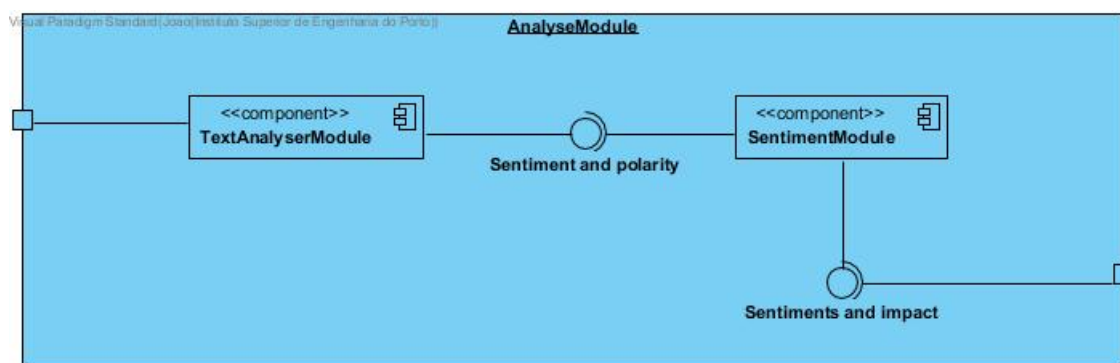


Figura 3.2: Modelo de componentes do módulo de análise emocional

3.2 Alternativa da arquitetura do sistema

Na secção 3.1 expomos e explicamos a arquitetura escolhida para a solução. Agora, nesta secção, iremos apresentar uma alternativa a essa arquitetura. A alternativa também seguirá o princípio *Single-Responsability Principle*.

O tratamento de texto seria feito por um módulo, que transformaria o texto fornecido numa árvore de *Kernel*. Logo, teria mais um módulo que a arquitetura anterior. Esse módulo trataria do texto e utilizaria o *lexicon WordNet* para ir buscar a emoção relativa a essa palavra. Os restantes módulos seriam iguais aos da proposta presente na 3.1. A arquitetura em análise é exposta na figura 3.3.

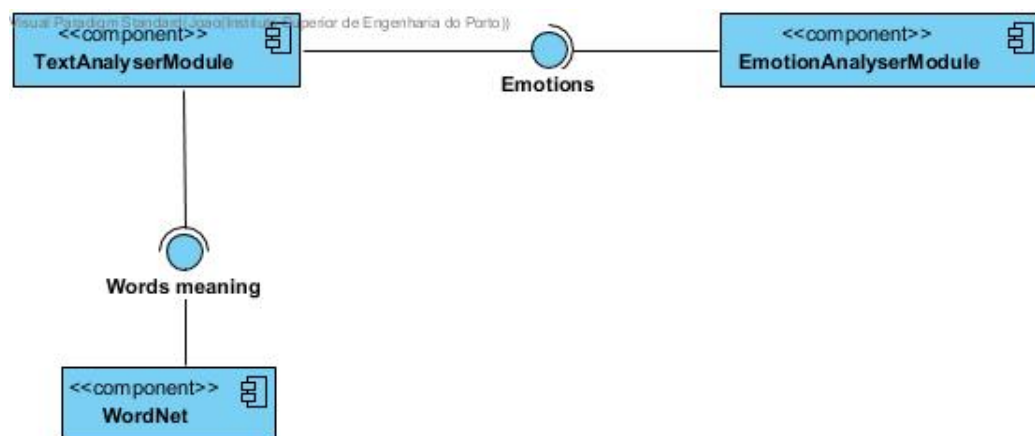


Figura 3.3: Modelo alternativo de componentes do sistema

3.3 Desenvolvimento da solução

Como explicado na secção 3.1, a solução será composta por um módulo que é dividido em dois: o *TextAnalyserModule* e o *EmotionAnalyserModule*. Nesta secção iremos explicar cada componente do protótipo desenvolvido, com base da pesquisa efetuada.

A escolha da linguagem computacional *python*, para este projeto, deveu-se a três fatores:

- A linguagem *python* tem uma curva de aprendizagem curta, ou seja, é simples aprender a sua utilização;
- A linguagem *python* tem várias ferramentas eficazes para *Natural language Processing*. A *nlTK* é das ferramentas mais completas e utilizadas;
- O *lexicon SenticNet* possui uma biblioteca compatível com *python*, facilitando a sua utilização.

A solução está dividida em dois módulos, o *TextAnalysisModule* e o *SentimentModule*. O módulo *TextAnalysisModule* é responsável pelo tratamento de texto. Este módulo transforma o ficheiro *Comma-separated values* (csv) em informação legível e envia-o para o *SentimentModule*. Por fim, o módulo *SentimentModule* gera um ficheiro. As emoções, o seu impacto e o seu número estão contidos nesse ficheiro, em formato csv.

3.3.1 TextAnalysisModule

Nesta subsecção iremos explicar o processo do módulo de análise de texto. Seguindo o fluxo da execução do módulo, a primeira ação é ler e carregar a informação de um *dataset*. A maioria dos *datasets* estão em ficheiros com o formato csv. Para ler esses ficheiros, é utilizada a biblioteca *pandas*. *Pandas* é uma biblioteca criada para a linguagem *python*, com a função de manipulação de dados. Para além de conter métodos para leitura e criação de ficheiros, permite criar estruturas de dados simples e rápidas de manipular. Estas estruturas são semelhantes a uma matriz ou tabela.

Os *datasets* utilizados neste projeto têm uma estrutura composta por uma ou várias colunas com texto para análise. O texto pode ser em forma de discurso ou frases simples.

Opcionalmente, os *datasets* podem conter uma coluna com a seguinte informação: se os intervenientes estão a concordar ou não um com o outro. Os textos devem estar na língua inglesa.

Na figura 3.4 está exposto um *dataset* com uma estrutura e conteúdo ideal para o bom funcionamento do módulo. Esse *dataset* é um conjunto de dados com argumentações entre duas pessoas e com a informação se concordam ou não entre si. O *dataset* foi fornecido pelo grupo de investigação GECAD. O último é composto por três colunas: a *arg1*, que representa o argumento do primeiro arguente; a *arg2*, que representa o argumento do segundo arguente e a *relation*, que representa se ambos os intervenientes concordam (*support*) ou se discordam (*attack*) um com o outro.

arg1	arg2	relation
This issue is about late-term abortions, in th	It is a bit of butchery.	attack
This issue is about late-term abortions, in th	If the doctor says he has to, I mean surgery is also k	support
It is a bit of butchery.	If the doctor says he has to, I mean surgery is also k	attack
This issue is about late-term abortions, in th	Rough one... Description's gruesome but does the	attack
This issue is about late-term abortions, in th	Why this is banned ? If the mother or the foetus is	support
Rough one... Description's gruesome but dcl	agree with participant4 ;)	support
This issue is about late-term abortions, in th	So the foetus is formed, but if it can live after the b	support
The topic of the second debate is about BA	We can't ban the abortion in general, it's a right for	attack
The topic of the second debate is about BA	In the case of malformed or severely sick foetus, I v	attack
In the case of malformed or severely sick fo	I said mother in general, but also father should hav	support
In the case of malformed or severely sick fo	do not agree with this: "Put yourself at the site of	attack
I do not agree with this: "Put yourself at the	But you can decide to do the abortion during the tl	attack
I do not agree with this: "Put yourself at the	I don't see it as cruelty but it is a gesture of compas	attack

Figura 3.4: Exemplo de um possível dataset

Após a leitura do ficheiro, as colunas do *dataset* são guardadas numa estrutura de dados, que mantém a estrutura do *dataset*. Da estrutura de dados resultante são selecionadas as colunas que contêm os argumentos. Caso a coluna com a informação de concordância esteja no *dataset*, será guardada numa estrutura à parte. De seguida, as colunas de texto do *dataset*, que estão presentes numa estrutura de dados, são separadas. Cada uma dessas colunas será guardada na sua respetiva estrutura de dados. Dependendo do número de colunas presentes no *dataset*, os números de estruturas de dados criados podem variar.

No exemplo do *dataset* exposto na figura 3.4, verifica-se que o algoritmo vai criar três estruturas de dados: uma estrutura de dados com o texto presente na coluna *arg1*; uma estrutura de dados com o texto presente na coluna *arg2*; e outra com a informação da coluna *relations*.

Para determinar as emoções presentes em cada argumento, utilizando o *lexicon SenticNet*, a API tem que receber palavras.

A biblioteca *nltk* foi utilizada com o objetivo de filtrar as palavras de cada frase e de cada argumento. O primeiro passo é dividir o argumento em frases e retirar todas as vírgulas, parágrafos, chavetas e apostrofes das mesmas. Neste passo são utilizadas duas funções: a *strip* e a *replace*. Funções já existentes em *python*. A função *strip* faz uma cópia de um conjunto de caracteres sem espaços em branco e sem caracteres que se pretendam

eliminar, que no caso desta dissertação são as chavetas retas. A função *replace* substitui todos os caracteres escolhidos por outros caracteres. No segundo passo, pretende-se isolar cada palavra e retirar as chamadas *stop words*, que são palavras que não têm peso na identificação das emoções. Essas *stop words* são dadas pela biblioteca *nltk*. A *nltk* tem um conjunto de expressões regulares e algoritmos preparados para obter palavras ou frases de língua inglesa. Através dessa funcionalidade obtém-se uma lista com todas as frases utilizadas num argumento, como explicado em 2.3. Após obter-se frases de cada argumento volta-se a utilizar a biblioteca *nltk* para retirar pontuações de cada uma das frases. As frases são guardadas em variáveis.

Cada frase filtrada é enviada para três métodos.

O primeiro método é o *getMoodtags*, que recebe a frase e a separa por palavras. Utiliza o *plugin do SenticNet* para comunicar com a API do *SenticNet* e recebe a emoção correspondente de cada palavra. Uma vez recebidas as emoções de cada palavra, estas são contadas. O resultado é uma frase com as emoções encontradas seguidas pelo número de vezes que aparecem. Um exemplo desse resultado é "#interest:2#admiration:1". As emoções que podem ser obtidas são: interesse, admiração, felicidade, surpresa, tristeza, repugnância, medo e raiva. O segundo método é o *getPolarity* que recebe uma frase, separa-a por palavras e utiliza o *plugin SenticNet* para determinar a polaridade dessas palavras. Depois são todas somadas e divididas pelo número de conceitos que apareceram, de modo a saber a polaridade da frase. Por último, o terceiro método *getSentics*, muito semelhante ao funcionamento do método da polaridade, recebe uma frase e é obtido o grau de *pleasantness*, *aptitude*, *attention* e *sensitivity* de cada palavra. Com a média de cada um dos graus, consegue-se determinar o grau de *pleasantness*, *aptitude*, *attention* e *sensitivity* para cada frase.

No final deste módulo é formada uma lista com a quantidade de vezes que cada uma das emoções aparece e outra lista com a polaridade e os graus de *pleasantness*, *aptitude*, *attention* e *sensitivity* de cada argumento.

3.3.2 SentimentModule

Nesta subsecção serão analisadas as emoções obtidas no módulo de *TextAnalysisModule*, de forma a obter o seu impacto e apresentar as emoções de forma organizada.

Após o cálculo das emoções presentes em cada frase, verifica-se qual delas é a mais predominante, de modo a ser possível determinar o impacto emocional. A maneira de se calcular a emoção mais predominante é através dos passos seguintes:

1. Verificar qual a emoção que tem maior número;
2. Caso o número de emoções seja igual é verificada a polaridade da frase, vendo se é positiva ou negativa, e o grau de *pleasantness*, *aptitude*, *attention* e *sensitivity*, pois cada emoção tem o seu grau numa das dimensões emocionais 2.4.2.

No estudo (G. van Kleef, C. De Dreu e A. Manstead 2010) são referidos os possíveis impactos de cada emoção dependendo do ambiente em que os intervenientes estão inseridos: ambiente competitivo ou cooperativo. O modelo utilizado nesse estudo, o modelo *EASI*, está exposto na tabela 2.8. O modelo *EASI* mostra se a cooperação ou competição aumentam em relação às emoções e ao ambiente que intervenientes estão inseridos.

Nos estudos (Van Kleef, De Dreu e Manstead 2006; G. A. van Kleef, C. K. W. De Dreu e A. S. R. Manstead 2004), o investigador prevê o impacto que diversas emoções têm ao serem utilizadas em discursos ou argumentos nos outros intervenientes da negociação. O impacto varia com as emoções utilizadas e com o grau de confiança que os intervenientes têm entre si. O autor refere também que num ambiente cooperativo existe um elevado grau de confiança e num ambiente competitivo, um baixo grau de confiança. Neste trabalho o impacto é o grau de exigência e o grau de cedência. O mesmo estudo foi mais detalhado na secção 2.4.5

A escolha destes estudos, para determinar o impacto emocional, deveu-se ao facto de serem trabalhos ainda atuais e completos que analisam as emoções muito semelhantes às emoções obtidas pela *framework SenticNet*.

Devido às particularidades desta dissertação, desenvolveu-se uma adaptação do modelo *EASI* com os estudos anteriores. A nossa adaptação alterou as emoções que estavam identificadas no modelo *EASI* para emoções equivalentes. Porém, caso não fosse claro o ambiente em que os intervenientes se encontravam, seria utilizado o estudo de *Bersade*. Se estivessem presentes emoções positivas (felicidade, surpresa, admiração e interesse) ou polaridade positiva, na tomada de decisão, seria mais fácil existir acordo. Se estivessem presentes emoções negativas (raiva, medo, tristeza e repugnância) seria mais complicado existir acordo.

O modelo *EASI* determina se o grau de cooperação e competitividade aumenta ou diminui. Na adaptação desenvolvida nesta dissertação, também foi acrescentado o grau de cedência e de exigência, baseada nos artigos (Van Kleef, De Dreu e Manstead 2006; G. A. van Kleef, C. K. W. De Dreu e A. S. R. Manstead 2004).

O modelo *EASI* classifica os ambientes como competitivos (baixa confiança) ou cooperativos (alta confiança). Nesta dissertação o módulo não sabe o grau de confiança, mas sim se os intervenientes concordam ou se discordam entre si. Após alguma pesquisa foi possível verificar que se alguém concordar com alguém há grande probabilidade de estarmos perante um ambiente de confiança, (Banerjee e Sanyal 2012). Logo, o que é entendido no modelo *EASI* por alta confiança pode ser equiparado à concordância e o que é entendido por baixa confiança pode ser equiparado à discordância.

Na implementação do modelo adaptado, são verificadas as emoções predominantes em cada argumento e se há concordância ou não entre os intervenientes. Há diferentes *outputs*, dependendo dessa concordância. Deve-se ter em consideração que as emoções da tabela são um pouco diferentes das do *hourglass of emotions*, que são as obtidas no módulo *TextAnalysisModule*. Também se procedeu a uma adaptação dos quatro grupos de emoções do modelo *EASI*: o grupo de felicidade, alegria e contentamento passou a identificar as emoções de interesse, de admiração, de felicidade e de surpresa; o grupo de raiva, frustração e irritação passou a ser identificado pelas emoções raiva e repugnância; a tristeza já é obtida no *output*; e o grupo emocional culpa corresponde a medo e repugnância. O modelo adaptado verifica-se na tabela 3.1.

Tabela 3.1: Impacto de diversas emoções em ambientes de concordância e discordância

Emoções	Ambientes de concordância				Ambientes de discordância			
	E	Coo	Cm	Cd	E	Coo	Cm	Cd
Emoções positivas	+	+	-	+	+	-	+	+
Raiva e repugnância	-	-	+	-	+	+	-	-
Tristeza	-	+	-	0	-	-	+	0
Medo e repugnância	+	+	-	0	+	-	+	0

A tabela 3.1 mostra o modelo de possíveis impactos emocionais em terceiros, cujos resultados obtidos, dependem se se está presente de um ambiente cooperativo ou de um ambiente competitivo. Os resultados são divididos em quatro critérios: o nível exigência, representado por E na tabela; o grau de cooperação e de competitividade, representado por Coo e Cm, respectivamente; e o nível de cedência, representado por Cd na tabela. Os resultados dos quatro critérios podem variar entre aumento (+), diminuição (-) ou não ser possível saber (0).

Os *outputs* do impacto de cada argumento são frases que dependem dos resultados do módulo exposto na tabela 3.1. Depois de se verificar se os intervenientes estão a concordar, ou não, para cada tipo de emoção é somada a quantidade de vezes que esta aparece em cada argumento. A emoção que aparecer mais vezes é a que vai ditar o resultado. As frases que correspondem ao impacto são as seguintes:

- Aumento da cooperação, aumento do nível de exigência e maior facilidade de cedência;
- Redução de exigência e de cooperação
- Aumento da cooperação e diminuição de nível de exigência;
- Aumenta o nível de exigência e redução de competição;
- O impacto é o aumento de exigência, com maior facilidade de um acordo;
- Redução da exigência e aumento de cooperação;
- Aumento da competição e diminuição do nível de exigência;
- Aumenta o nível de exigência e aumenta a competição;
- O impacto é o aumento da probabilidade de cooperação;
- O impacto é o redução da probabilidade de cooperação.

Portanto se se estiver num ambiente de concordância e se as emoções predominantes forem emoções positivas, teremos um aumento da cooperação, um aumento do nível de exigência e uma maior facilidade de cedência. Se as emoções predominantes forem a raiva e a repugnância, então, teremos uma redução de exigência e de cooperação. Se a emoção for tristeza teremos um aumento da cooperação e uma diminuição do nível de exigência. Se a emoção predominante for medo e repugnância, teremos um aumento do nível de exigência e uma redução do nível de competição.

Se estivermos num ambiente de discordância, tendo presentes emoções positivas, teremos um aumento da competição, um aumento do nível de exigência e uma maior facilidade de cedência. Se as emoções predominantes forem raiva e repugnância teremos uma redução

da exigência e um aumento de cooperação. Se a emoção predominante for a tristeza teremos um aumento da competição e uma diminuição do nível de exigência. Por fim, se predominarem as emoções medo e repugnância, então, teremos um aumento do nível de exigência e um aumento da competição. Esta informação será guardada numa lista.

O resultado do módulo será guardado numa estrutura de dados do tipo *pandas* com as seguintes colunas:

- Column1 - Tem um identificador, que é um valor inteiro único, para cada argumento;
- IdDiscurso - Que serve para nos informar a que discurso pertence o argumento;
- Discurso - Representa o argumento que foi analisado;
- interest - Número de vezes que a emoção interesse aparece no argumento;
- admiration - Número de vezes que a emoção “admiração” aparece no argumento;
- joy - Número de vezes que a emoção “felicidade” aparece no argumento;
- surprise - Número de vezes que a emoção “surpresa” aparece no argumento;
- sadness - Número de vezes que a emoção “tristeza” aparece no argumento;
- disgust - Número de vezes que a emoção “repugnância” aparece no argumento;
- fear - Número de vezes que a emoção “medo” aparece no argumento;
- anger - Número de vezes que a emoção “raiva” aparece no argumento;
- Polarity - Valor da polaridade do argumento;
- Sentics - O grau de *pleasantness*, *aptitude*, *attention* e *sensitivity* do argumento;
- Impact - Representa o impacto que as emoções presentes no argumento têm nos restantes intervenientes da tomada de decisões.
- E - Representa o grau de exigência dos restantes intervenientes na tomada de decisão, que pode diminuir (-) ou aumentar (+);
- Co - Representa o grau de cooperação dos restantes intervenientes da tomada de decisão, que pode diminuir (-) ou aumentar (+);
- Cm - Representa o grau de competição dos restantes intervenientes na tomada de decisão, que pode diminuir (-) ou aumentar (+);
- Cd - Representa o grau de cedência dos restantes intervenientes na tomada de decisão, que pode diminuir (-), aumentar (+) ou não ser possível determinar (0);

Por fim, a estrutura de dados é guardada num ficheiro de formato *csv*.

Capítulo 4

Avaliação e experimentação

Neste capítulo iremos referir como a solução foi avaliada, bem como as metodologias utilizadas na sua experimentação. Na secção 4.1 são apresentadas as métricas de avaliação e hipóteses utilizadas. Na secção 4.2 vão ser apresentadas as metodologias utilizadas para avaliar da melhor maneira a solução. Na secção 4.3 indicam-se os tipos de testes utilizados e a sua caracterização.

4.1 Métricas de avaliação e hipóteses

De modo a avaliar algoritmos ou módulos de algoritmos tem de se determinar as melhores métricas, levantar hipóteses e analisar se essas hipóteses se verificam.

As métricas de avaliação utilizadas, para os módulos desta dissertação, interligam-se com complexidade e eficácia para determinar as emoções num texto e o impacto das mesmas. A métrica escolhida para compreender o desempenho do módulo é o *precision*, *recall* e *F-score*. A precisão ou *precision* (P) é a razão do que foi considerado verdadeiro que é realmente verdadeiro (TP), entre o que foi considerado como verdadeiro (TP+FP), presente na equação 4.1.

$$P = \frac{TP}{TP + FP} \quad (4.1)$$

O *recall* (R) é a razão do que foi considerado verdadeiro que é realmente verdadeiro (TP), entre o que realmente é verdadeiro (TP+FN), presente na equação 4.2.

$$R = \frac{TP}{TP + FN} \quad (4.2)$$

O *F-score* é algo que relaciona a precisão com o *recall*. A equação que determina o *F-score* está presente na 4.3

$$Fs = 2 * \frac{precision * recall}{precision + recall} \quad (4.3)$$

Como já foi referido, para que os algoritmos sejam devidamente avaliados, é necessário o levantamento de hipóteses. As hipóteses podem ser alteradas ou até acrescentadas no decorrer da análise. As hipóteses a serem confirmadas são:

- É possível determinar as emoções presentes num dado texto;
- É possível determinar o impacto emocional das emoções presentes no texto.

4.2 Metodologias de avaliação

Nesta subsecção, vai ser explicada a metodologia escolhida para avaliar se os algoritmos conseguiram concretizar as hipóteses. A avaliação consiste em aplicar simulações de validação de hipóteses.

Os testes servem para verificar a eficiência do código dos algoritmos. Os testes poderão detetar erros mesmo antes de executar as simulações. Este processo poupa tempo e é mais preciso. Os tipos de testes utilizados vão ser abordados na secção 4.3.

As simulações ajudarão a entender se o sistema desenvolvido estará a fazer a análise de emoções, corretamente. As simulações executam o sistema com os dados fornecidos de um *dataset*. Após a execução, será verificada a eficácia de cada algoritmo utilizando as métricas referidas na secção 4.1.

Nesta dissertação serão efetuadas duas simulações, uma para avaliar a funcionalidade de detetar emoções e outra para avaliar a determinação do impacto das mesmas.

4.2.1 Experiência com a deteção de emoções

A experiência efetuada para avaliar a deteção de emoções passou por utilizar os dados de um *dataset*, que foram, por sua vez, utilizados no módulo desenvolvido nesta dissertação. O *dataset* escolhido é um excerto do *dataset ISEAR*. O *dataset ISEAR* (*international survey on emotions antecedents and reactions dataset*) foi desenvolvido pela Universidade de Genebra. O último *dataset* contém dados relacionados com sete emoções (felicidade, medo, raiva, tristeza, repugnância, vergonha e culpa), que foram retiradas de questionários a alunos de todo o mundo. Pelo facto de cada argumento, dos questionários, conter emoções predominantes já esperadas e por essas emoções serem semelhantes às obtidas no nosso módulo, achamos interessante escolher o *dataset* da Universidade de Genebra, para desenvolver a nossa experiência. Porém, não utilizamos o mesmo *dataset* que utilizamos no desenvolvimento da solução.

O *dataset* contém duas colunas: *content* e *sentiment*. A coluna *content* contém as frases a serem analisadas, todas em língua inglesa. A coluna *sentiment* tem a emoção mais predominante de cada uma das frases da coluna *content*. As emoções presentes são: raiva, tristeza, medo e felicidade. Conforme já dissemos, a escolha do *dataset* deveu-se a este último conter já as emoções predominantes de cada frase, de modo a facilitar a comparação com os resultados do módulo. E também por as quatro emoções presentes na *dataset* pertencerem aos possíveis resultados do módulo.

Após o módulo ser executado, utilizando os dados do *dataset*, foi recolhido um resultado, que foi comparado com o resultado esperado. Nessa comparação foram utilizadas as métricas *precision*, *recall* e *F-score* para cada conjunto de emoções. No caso das frases que tinham como emoção predominante a raiva foi:

- Calculada a precisão através da divisão do número de vezes que a emoção “raiva” foi corretamente identificada pelo número de vezes que a emoção “raiva” foi identificada;
- Calculado o *recall* através da divisão do número de vezes que a emoção “raiva” foi corretamente identificada pelo número de vezes que a emoção “raiva” deveria ter sido identificada;

- Calculado o *F-score* através da equação 4.1.

Os passos enumerados anteriormente foram aplicados também nas emoções “tristeza”, “felicidade” e “medo”. À raiva foi associada a repugnância, que o módulo também determina, por serem muito semelhantes. Para a emoção “raiva” o módulo apresentou uma precisão de 0.3971, *recall* de 0.4991 e um *F-score* de 0.4423. Na emoção “felicidade” o módulo apresentou uma precisão de 0.445, *recall* de 0.6578 e um *F-score* de 0.5309. Na emoção “tristeza” o módulo apresentou uma precisão de 0.4641, *recall* de 0.1474 e um *F-score* de 0.2237. Na emoção “medo” o módulo apresentou uma precisão de 0.4037, *recall* de 0.3743 e um *F-score* de 0.3884. Os valores desta experiência foram comparados aos trabalhos de Agrawal e An (2012) e Gievska, Koroveshovski e Chavdarova (2014), como se verifica na tabela 4.1.

Tabela 4.1: Resultados dos *F-scores* de vários módulos

Módulos	Raiva	Felicidade	Tristeza	Medo	Média
keyword baseline	0.346	0.371	0.270	0.328	0.328
DIM	0.286	0.515	0.337	0.351	0.372
Hybrid method	0.899	0.803	0.877	0.903	0.84
Módulo SenticNet	0.442	0.531	0.224	0.388	0.396

Através das comparações das experiências efetuadas nesta dissertação e das experiências nos artigos (Agrawal e An 2012; Gievska, Koroveshovski e Chavdarova 2014), concluiu-se que este trabalho tem um bom desempenho comparado a outros que só utilizam *lexicons* ou métodos de *machine learning*. Porém, os métodos híbridos, módulos que utilizam *lexicons* em conjunto com métodos de *machine learning*, já têm um desempenho muito superior ao do módulo *SenticNet* desenvolvido nesta dissertação.

Numa análise mais profunda, aos resultados, verifica-se que o módulo tem um melhor desempenho quando se está a identificar emoções de felicidade e de raiva. Uma razão que influencia o resultado do módulo é que este identifica oito emoções, ao contrário das quatro, que estão identificadas no *dataset*. Isto causa alterações ao cálculo da polaridade e dos graus de *pleasantness*, *attention*, *sensitivity* e *aptitude*, que por sua vez vai alterar a emoção predominante.

4.2.2 Experiência com impacto emocional

A experiência efetuada para a avaliação do impacto emocional consistiu na utilização do *dataset* referido na secção 3.3.1. Este *dataset* foi escolhido por ter a estrutura indicada para o funcionamento integral do módulo desenvolvido, nesta dissertação. O conjunto de dados contém 263 argumentações. Cada argumentação tem dois argumentos e a concordância entre os dois arguentes.

Os resultados desta experiência podem ser observados nas figuras 4.1, 4.2, 4.4, 4.5 e 4.3. A figura 4.1 possui um excerto do resultado do módulo. A identificação dos argumentos e dos discursos, encontra-se na coluna *IdDiscurso*; na coluna *Discurso* estão os argumentos; e na coluna *Column1* estão os identificadores.

Column1	IdDiscurso	Discurso
0	Argumento 1 do discurso 1	['This issue is about late-term abortions, in the second and third trimesters.', 'At least 90% of a
1	Argumento 1 do discurso 2	['This issue is about late-term abortions, in the second and third trimesters.', 'At least 90% of a
2	Argumento 1 do discurso 3	['It is a bit of butchery.']
3	Argumento 1 do discurso 4	['This issue is about late-term abortions, in the second and third trimesters.', 'At least 90% of a
4	Argumento 1 do discurso 5	['This issue is about late-term abortions, in the second and third trimesters.', 'At least 90% of a
5	Argumento 1 do discurso 6	['Rough one... ', 'Description's gruesome but does the foetus fully lives at that point and therefo
6	Argumento 1 do discurso 7	['This issue is about late-term abortions, in the second and third trimesters.', 'At least 90% of a
7	Argumento 1 do discurso 8	['The topic of the second debate is about BANNING PARTIAL BIRTH ABORTIONS.']
8	Argumento 1 do discurso 9	['The topic of the second debate is about BANNING PARTIAL BIRTH ABORTIONS.']
9	Argumento 1 do discurso 10	['In the case of malformed or severely sick foetus, I would definitely understand the need to me
10	Argumento 1 do discurso 11	['In the case of malformed or severely sick foetus, I would definitely understand the need to me
11	Argumento 1 do discurso 12	['I do not agree with this: "Put yourself at the site of the diseased or malformed children makes
12	Argumento 1 do discurso 13	['I do not agree with this: "Put yourself at the site of the diseased or malformed children makes
13	Argumento 1 do discurso 14	['I do not agree with this: "Put yourself at the site of the diseased or malformed children makes
14	Argumento 1 do discurso 15	['I don't see it as cruelty but it is a gesture of compassion for the unborn in the case of necessi
15	Argumento 1 do discurso 16	['Think about the possible baby after birth, if it's malformed or have an important disease it cai
16	Argumento 1 do discurso 17	['If you know the future baby will be malformed or otherwise, it's a crime not to do something f
17	Argumento 1 do discurso 18	['It's not selection because some babies with diseases are born, but a baby who can't live by h
18	Argumento 1 do discurso 19	['It's not selection because some babies with diseases are born, but a baby who can't live by h
19	Argumento 1 do discurso 20	['The topic of the second debate is about BANNING PARTIAL BIRTH ABORTIONS.']

Figura 4.1: Exemplo de um possível ficheiro resultado do algoritmo das colunas: Column1, IdDiscurso e Discurso

A figura 4.2 expõe um excerto dos resultados das emoções encontradas. Em cada coluna mostra-se o número de vezes que cada uma das emoções é identificada nas frases.

interest	admiration	joy	surprise	sadness	disgust	fear	anger
17	27	30	6	15	13	6	2
17	27	30	6	15	13	6	2
0	1	1	0	0	1	0	1
17	27	30	6	15	13	6	2
17	27	30	6	15	13	6	2
5	9	7	1	1	4	0	3
17	27	30	6	15	13	6	2
0	0	2	2	0	0	0	0
0	0	2	2	0	0	0	0
5	8	7	0	5	4	3	4
5	8	7	0	5	4	3	4
2	5	5	0	2	2	3	1
2	5	5	0	2	2	3	1
2	5	5	0	2	2	3	1
7	10	2	1	1	1	1	1
4	5	3	0	2	2	1	1
1	3	0	0	2	1	2	1
4	5	3	0	1	0	1	0
4	5	3	0	1	0	1	0
0	0	2	2	0	0	0	0

Figura 4.2: Exemplo de um possível ficheiro resultado do algoritmo das colunas: Interest, admiration, joy, surprise, sadness, disgust, fear, anger

A figura 4.3 expõe um excerto dos resultados das polaridades e dos graus de *pleasantness*, *aptitude*, *attention* e *sensitivity* de cada um dos argumentos.

Polarity	Sentics
0.1621896551724138	{'pleasantness': 0.01273684167737419, 'attention': 0.015060496552212714, 'sensitivity': -8.365858}
0.1621896551724138	{'pleasantness': 0.01273684167737419, 'attention': 0.015060496552212714, 'sensitivity': -8.365858}
-0.0310000000000000028	{'pleasantness': -0.008000000000000007, 'attention': -0.285, 'sensitivity': 0.469, 'aptitude': 0.09900}
0.2008793103448276	{'pleasantness': 0.01273684167737419, 'attention': 0.015060496552212714, 'sensitivity': -8.365858}
0.2008793103448276	{'pleasantness': 0.01273684167737419, 'attention': 0.015060496552212714, 'sensitivity': -8.365858}
0.2236666666666667	{'pleasantness': 0.004215246707440242, 'attention': 0.010007651227798632, 'sensitivity': 2.548084}
0.2008793103448276	{'pleasantness': 0.01273684167737419, 'attention': 0.015060496552212714, 'sensitivity': -8.365858}
0.4425	{'pleasantness': 0.47550000000000003, 'attention': -0.44, 'sensitivity': 0.0, 'aptitude': 0.389}
0.4425	{'pleasantness': 0.47550000000000003, 'attention': -0.44, 'sensitivity': 0.0, 'aptitude': 0.389}
-0.010333333333333318	{'pleasantness': 0.003440172220144116, 'attention': 0.0001836335160878849, 'sensitivity': -0.0416}
-0.010333333333333318	{'pleasantness': 0.003440172220144116, 'attention': 0.0001836335160878849, 'sensitivity': -0.0416}
-0.019799999999999974	{'pleasantness': 0.0006273726851851851, 'attention': 0.0014412020502645503, 'sensitivity': -0.086}
-0.019799999999999974	{'pleasantness': 0.0006273726851851851, 'attention': 0.0014412020502645503, 'sensitivity': -0.086}
-0.019799999999999974	{'pleasantness': 0.0006273726851851851, 'attention': 0.0014412020502645503, 'sensitivity': -0.086}
0.1435833333333333	{'pleasantness': -0.005024035623263054, 'attention': 0.03227393574676995, 'sensitivity': -0.00849}
0.3105555555555553	{'pleasantness': 0.07540437610229277, 'attention': 0.0948054756393298, 'sensitivity': -0.01114945}
-0.28159999999999996	{'pleasantness': -0.19418333333333332, 'attention': 0.001375, 'sensitivity': -0.12195, 'aptitude': -0.0}

Figura 4.3: Exemplo de um possível ficheiro resultado do algoritmo das colunas: Polarity e Sentics

A figura 4.4 tem um excerto dos resultados das frases com dos possíveis impactos dos argumentos.

Impact
O impacto é o aumento de exigência, porém com maior facilidade de um acordo.
Aumento da cooperação, aumento do nível de exigência e maior facilidade de cedência
Redução da exigência e aumento de cooperação.
O impacto é o aumento de exigência, porém com maior facilidade de um acordo.
Aumento da cooperação, aumento do nível de exigência e maior facilidade de cedência
Aumento da cooperação, aumento do nível de exigência e maior facilidade de cedência
Aumento da cooperação, aumento do nível de exigência e maior facilidade de cedência
O impacto é o aumento de exigência, porém com maior facilidade de um acordo.
O impacto é o aumento de exigência, porém com maior facilidade de um acordo.
Aumenta o nível de exigência e redução de competição.
Aumenta o nível de exigência e aumenta a competição.
Aumento da competição e diminui o nível de exigência.
Aumento da competição e diminui o nível de exigência.
Aumento da competição e diminui o nível de exigência.
Aumento da cooperação, aumento do nível de exigência e maior facilidade de cedência
Aumento da cooperação, aumento do nível de exigência e maior facilidade de cedência
Aumento da cooperação e diminuição de nível de exigência.
Aumento da cooperação, aumento do nível de exigência e maior facilidade de cedência

Figura 4.4: Exemplo de um possível ficheiro resultado do algoritmo da coluna Impact

A figura 4.5 expõe um excerto dos resultados do nível de exigência, do nível da probabilidade de cooperação, do nível da probabilidade de competição e do nível de cedência.

Nível de exigência ▼	Probabilidade de cooperação ▼	Probabilidade de competição ▼	Nível de cedência ▼
+	-	+	+
+	+	-	+
+	+	-	-
+	-	+	+
+	+	-	+
+	+	-	+
+	+	-	+
+	-	+	+
+	-	+	+
+	+	-	0
+	-	+	0
-	-	+	0
-	-	+	0
-	-	+	0
+	+	-	+
+	+	-	+
-	+	-	0
+	+	-	+

Figura 4.5: Exemplo de um possível ficheiro resultado do algoritmo das colunas: E, Co, Cm e Cd

Nesta experiência não foi possível aplicar as métricas de avaliação, porque não se conhecia

o resultado esperado. O resultado que se pretendia era o impacto emocional de cada argumento que foi sentido pelos restantes intervenientes. Não foi encontrado outro *dataset* que tivesse o impacto esperado. Logo, não foi possível avaliar esta experiência.

4.3 Tipos de testes utilizados

Como já referido, vários testes foram desenvolvidos para testar a eficiência e a qualidade dos algoritmos dos módulos da solução. Os tipos de testes utilizados foram:

- Testes unitários – Este tipo de teste verifica cada algoritmo, de forma a testar a validade dos mesmos. Para que seja possível, vai-se introduzir *inputs* nos algoritmos, para verificar se cada algoritmo dá o resultado esperado.
- Testes de desempenho – Os testes de desempenho servem para verificar qual o melhor algoritmo dos estudados. Ou seja, através deste tipo de teste verifica-se o desempenho de um tipo de algoritmo para depois ser comparado com os outros, visando escolher o melhor. O resultado vai determinar os algoritmos a serem usados, utilizando a métrica definida na secção 4.1.

Capítulo 5

Conclusões

O capítulo presente apresenta as conclusões desta dissertação. Na secção 5.1 será apresentado um pequeno resumo do que foi conseguido com esta dissertação e uma análise do cumprimento dos objetivos e hipóteses. A secção 5.2 irá expor as limitações do projeto. Por fim, a secção 5.3 descreve propostas de possíveis ideias para trabalhos futuros.

5.1 Síntese do projeto

As emoções estão presentes em qualquer tomada de decisão e podem influenciá-la, mesmo que os intervenientes não se apercebam. Numa tomada de decisão em grupo, cada interveniente tenta convencer ou expor a sua opinião. Ao argumentar, numa tomada de decisão, as emoções presentes em cada argumento podem, de certa forma, influenciar os restantes intervenientes. A esse processo chama-se contágio de emoções. Um decisor sente emoções que são percebidas pelos restantes participantes, no processo de decisão. Caso haja um contágio, os restantes irão ser influenciados na sua decisão. A reação é aquilo que foi chamado de impacto emocional, nesta dissertação. Cada emoção tem o seu possível impacto. A pesquisa levou a que se verificasse uma escassez de trabalho no que toca à determinação do impacto emocional no âmbito de ciências computacionais na literatura. O módulo criado nesta dissertação tem como objetivos:

- Detetar e identificar automaticamente emoções em mensagens trocadas pelos decisores, em contexto de tomada de decisão em grupo;
- Determinar o impacto emocional de uma determinada afirmação nos restantes elementos participantes, num processo de tomada de decisão em grupo.

O capítulo do Estado da Arte serviu para expor e explicar todos os conceitos importantes para que fosse possível o desenvolvimento da dissertação. O tema principal desta dissertação foi computação afetiva e processamento de linguagem natural. No processamento de linguagem natural foi determinado que a ferramenta para tratamento de texto, mais apropriada, seria a *natural language toolkit*. Na secção da computação afetiva foram analisados conceitos da análise de personalidade, da análise de emoções, *datasets* e *lexicons* sobre análise emocional, abordagens para determinar e classificar emoções e o impacto emocional em terceiros. Foi concluído que o melhor *lexicon* era o *SenticNet*, por conter um dicionário compatível com muitas línguas e por analisar conceitos de uma maneira mais completa que os restantes. Com base no estudo que foi feito no Estado da Arte, o melhor algoritmo para análise emocional é o *Naive Bayes classifier*, na maioria das situações. No impacto emocional em terceiros, o modelo mais completo foi o modelo *EASI* de acordo com esse mesmo estudo.

No capítulo de design e implementação da solução foi descrita a arquitetura do módulo, bem como o seu desenvolvimento. O módulo referido é um conjunto de algoritmos que resulta nas emoções presentes no texto e o impacto que essas emoções podem ter causado aos restantes intervenientes na tomada de decisões ou argumentação.

A dissertação teve como principais contributos os pontos seguintes:

- Pesquisa e identificação de informação sobre emoções num contexto de sistemas de apoio à decisão em grupo;
- Pesquisa e identificação de informação sobre a extração do impacto de certas emoções de um texto, verificando a concordância ou não dos intervenientes;
- Design e implementação de um módulo que permite determinar as emoções presentes em argumentos e calcular o impacto das mesmas, num contexto de tomada de decisão.

No capítulo 4, demonstramos o tipo de métricas, as hipóteses que esta dissertação tinha de cumprir, os tipos de teste utilizados e explicamos e analisamos as experiências efetuadas. A avaliação e a experimentação, com a funcionalidade de determinação e deteção de emoções, foi um sucesso. Com essa avaliação concluímos que o módulo desenvolvido foi mais preciso do que outros, que utilizavam abordagens de *lexicon*. Mas é menos preciso do que as abordagens híbridas.

Verificamos que não foi possível avaliar a funcionalidade do cálculo de impacto emocional, mas foi possível experimentá-lo. Em suma, as hipóteses foram verificadas e os objetivos cumpridos.

5.2 Limitações do projeto

Nesta secção são exploradas as limitações do projeto elaborado nesta dissertação. As limitações foram as seguintes:

- O módulo desenvolvido não reconhece ironia nem sarcasmo, o que afeta a identificação das emoções presentes no discurso;
- O módulo tem como pressuposto o contágio de emoções. Porém, esse contágio pode não ocorrer, principalmente, quando estão presentes emoções positivas no argumento. As emoções positivas podem alterar o impacto emocional;
- Não foi possível avaliar corretamente o módulo que determina o impacto emocional.

5.3 Trabalhos futuros

Nesta dissertação foram analisados e pesquisados os temas de computação afetiva, de sistemas de apoio à decisão em grupo e de processamento de linguagem natural. Contudo, verifica-se que há ainda muito a fazer e a pesquisar. No sentido de melhorar e de combater as limitações que ainda persistem. De seguida, são apontadas algumas ideias para futuro desenvolvimento:

- O trabalho foi elaborado no contexto de sistemas de apoio à tomada de decisão em grupo. Mas, ainda não foi possível integrar o trabalho desenvolvido num desses

sistemas. A integração do módulo com um sistema de apoio à tomada de decisão em grupo é considerada um projeto importante. Tal como foi verificado nesta dissertação, as emoções têm bastante impacto numa decisão;

- A ironia e o sarcasmo são limitações do módulo. Uma das melhorias a ser executada passa pela sua deteção e pela verificação da sua influência no resultado do impacto emocional;
- O módulo desenvolvido só analisa texto. Por isso, um dos trabalhos a considerar seria acrescentar as funcionalidades de análise de áudio e de vídeo;
- Na pesquisa efetuada, concluiu-se que as metodologias para identificar emoções mais eficientes são as híbridas, utilizando *lexicon* e algoritmos de *machine learning*. Neste sentido, um dos trabalhos futuros pode passar pela incorporação de outro algoritmo de *machine learning* para auxiliar o já criado;
- Utilização de um sistema de multiagentes para simular uma argumentação e comparar os impactos, que os agentes tiveram, com os impactos determinados pelo módulo desenvolvido nesta dissertação.

Bibliografia

- Agarwal, Apoorv et al. (2011). «Sentiment Analysis of Twitter Data». Em: p. 9.
- Aggarwal, Charu C. e ChengXiang Zhai (2012). *Mining Text Data*. Springer Science & Business Media. 527 pp. isbn: 978-1-4614-3223-4. Google Books: vFH0x8wfSU0C.
- Agrawal, Ameeta e Aijun An (2012). «Unsupervised Emotion Detection from Text Using Semantic and Syntactic Relations». Em: *2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology*. 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology. Vol. 1, pp. 346–353. doi: 10.1109/WI-IAT.2012.170.
- Allport, Gordon W. (1961). *Pattern and Growth in Personality*. Pattern and Growth in Personality. Oxford, England: Holt, Reinhart & Winston. xiv, 593.
- Banerjee, Amitav e Debmitra Sanyal (2012). «Dynamics of Doctor–Patient Relationship: A Cross-Sectional Study on Concordance, Trust, and Patient Enablement». Em: *Journal of Family and Community Medicine* 19.1, pp. 12–19. issn: 2230-8229. doi: 10.4103/2230-8229.94006. pmid: 22518353. url: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3326765/> (acedido em 29/07/2020).
- Barbosa, Luciano e Junlan Feng (2010). «Robust Sentiment Detection on Twitter from Biased and Noisy Data». Em: p. 9.
- Belliveau, Paul, Abbie Griffin e Stephen Somermeyer (2004). *The PDMA ToolBook 1 for New Product Development*. John Wiley & Sons. 5-29. isbn: 978-0-471-27108-6. Google Books: kqX5EvT2U8AC.
- Bersade, Sigal G. (2002). *The Ripple Effect: Emotional Contagion and Its Influence on Group Behavior* - Sigal G. Barsade, 2002. url: <https://journals.sagepub.com/doi/abs/10.2307/3094912> (acedido em 16/06/2020).
- Bird, Steven, Ewan Klein e Edward Loper (2009). *Natural Language Processing with Python: Analyzing Text with the Natural ...* - Steven Bird, Ewan Klein, Edward Loper - Google Livros. 1ª. O'Reilly. url: https://books.google.pt/books?hl=pt-PT&lr=&id=KGibfiiP1i4C&oi=fnd&pg=PR5&dq=Natural+Language+Processing+with+Python&ots=Y3HiB1GEN0&sig=1MMspIxp_m_DJMFPLPTnP9_SfRg&redir_esc=y#v=onepage&q=Natural%20Language%20Processing%20with%20Python&f=false (acedido em 30/06/2020).
- Buehler, Roger, Deanna Messervey e Dale Griffin (2005). «Collaborative Planning and Prediction: Does Group Discussion Affect Optimistic Biases in Time Estimation?» Em: *Organizational Behavior and Human Decision Processes* 97.1, pp. 47–63. issn: 0749-5978. doi: 10.1016/j.obhdp.2005.02.004. url: <http://www.sciencedirect.com/science/article/pii/S0749597805000348> (acedido em 06/02/2020).
- Cambria, Erik, Andrew Livingstone et al. (2012). «The Hourglass of Emotions». Em: *Cognitive Behavioural Systems*. Ed. por Anna Esposito et al. Vol. 7403. Berlin, Heidelberg: Springer Berlin Heidelberg, pp. 144–157. isbn: 978-3-642-34583-8 978-3-642-34584-5. doi: 10.1007/978-3-642-34584-5_11. url: http://link.springer.com/10.1007/978-3-642-34584-5_11 (acedido em 29/01/2020).

- Cambria, Erik, Daniel Olsher e Dheeraj Rajagopal (2014). «SenticNet 3: A Common and Common-Sense Knowledge Base for Cognition-Driven Sentiment Analysis». Em: *Twenty-Eighth AAAI Conference on Artificial Intelligence*. Twenty-Eighth AAAI Conference on Artificial Intelligence. url: <https://www.aaai.org/ocs/index.php/AAAI/AAAI14/paper/view/8479> (acedido em 12/02/2020).
- Cambria, Erik, Soujanya Poria et al. (2018). «SenticNet 5: Discovering Conceptual Primitives for Sentiment Analysis by Means of Context Embeddings». Em: *Thirty-Second AAAI Conference on Artificial Intelligence*. Thirty-Second AAAI Conference on Artificial Intelligence. url: <https://www.aaai.org/ocs/index.php/AAAI/AAAI18/paper/view/16839> (acedido em 12/02/2020).
- Carneiro, João Miguel Ribeiro (2018). «Negotiation in group decision support systems - An approach based in argumentation and satisfaction». Em: url: <http://repositorium.sdum.uminho.pt/> (acedido em 18/01/2020).
- Chen, Chen-Tung (2000a). «Extensions of the TOPSIS for Group Decision-Making under Fuzzy Environment». Em: *Fuzzy Sets and Systems. An International Journal in Information Science and Engineering* 114.1, pp. 1–9. issn: 0165-0114. doi: 10.1016/S0165-0114(97)00377-1. url: <http://www.sciencedirect.com/science/article/pii/S0165011497003771> (acedido em 11/12/2019).
- (2000b). «Extensions of the TOPSIS for Group Decision-Making under Fuzzy Environment». Em: *Fuzzy Sets and Systems. An International Journal in Information Science and Engineering* 114.1, pp. 1–9. issn: 01650114. doi: 10.1016/S0165-0114(97)00377-1. url: <https://linkinghub.elsevier.com/retrieve/pii/S0165011497003771> (acedido em 06/02/2020).
- Eysenck, H. J. (1991). «Dimensions of Personality: 16, 5 or 3?—Criteria for a Taxonomic Paradigm». Em: *Personality and Individual Differences* 12.8, pp. 773–790. issn: 0191-8869. doi: 10.1016/0191-8869(91)90144-Z. url: <http://www.sciencedirect.com/science/article/pii/019188699190144Z> (acedido em 28/01/2020).
- Freitas, Carlos Filipe et al. (2013). «A Multi-Agents Framework for Contextual and Affective Decision Making». Em: *Intelligent Environments*. doi: 10.3233/978-1-61499-286-8-347.
- Gievska, Sonja, Kiril Koroveshovski e Tatjana Chavdarova (2014). «A Hybrid Approach for Emotion Detection in Support of Affective Interaction». Em: *2014 IEEE International Conference on Data Mining Workshop*. 2014 IEEE International Conference on Data Mining Workshop, pp. 352–359. doi: 10.1109/ICDMW.2014.130.
- Gross, J.J. e Ross Thompson (2007). «Emotion Regulation: Conceptual Foundations». Em: *Handbook of Emotion Regulation*, pp. 3–27.
- Hilgard, Ernest R. (1953). *Introduction to Psychology*. Introduction to Psychology. Oxford, England: Harcourt, Brace. x, 659.
- Hogenboom, Alexander et al. (2014). «Multi-Lingual Support for Lexicon-Based Sentiment Analysis Guided by Semantics». Em: *Decision Support Systems* 62, pp. 43–53. issn: 0167-9236. doi: 10.1016/j.dss.2014.03.004. url: <http://www.sciencedirect.com/science/article/pii/S0167923614000645> (acedido em 30/12/2019).
- Howard, Pierce J. e Jane M. Howard (1995). *The Big Five Quickstart: An Introduction to the Five-Factor Model of Personality for Human Resource Professionals*. url: <https://eric.ed.gov/?id=ED384754> (acedido em 21/01/2020).
- Kaner, Sam (2014). *Facilitator's Guide to Participatory Decision-Making - Sam Kaner - Google Livros*. url: <https://books.google.pt/books?hl=pt-PT&lr=&id=q-gRBAAQBAJ&oi=fnd&pg=PA20&dq=Facilitator%E2%80%99sguidetoparticipatorydecision->

- making&ots=eW4PES_Sua&sig=M3yD35Myj09246jByjij2V2ImE&redir_esc=y#v=onepage&q&f=false (acedido em 07/02/2020).
- Lerner, Jennifer S (2004). «Emotions and Decision Making, p.» Em: p. 45.
- Lin, Wei-Hao et al. (2006). «Which Side Are You on?: Identifying Perspectives at the Document and Sentence Levels». Em: *Proceedings of the Tenth Conference on Computational Natural Language Learning - CoNLL-X '06*. The Tenth Conference. New York City, New York: Association for Computational Linguistics, p. 109. doi: 10.3115/1596276.1596297. url: <http://portal.acm.org/citation.cfm?doid=1596276.1596297> (acedido em 29/01/2020).
- Loper, Edward e Steven Bird (2002). *NLTK: The Natural Language Toolkit*. arXiv: cs/0205028. url: <http://arxiv.org/abs/cs/0205028> (acedido em 02/10/2020).
- Loureiro, Daniel, Goreti Marreiros e José Neves (2011). «Sentiment Analysis of News Titles». Em: *Progress in Artificial Intelligence*. Ed. por Luis Antunes e H. Sofia Pinto. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 1–14. isbn: 978-3-642-24769-9. doi: 10.1007/978-3-642-24769-9_1.
- Luthans, Fred, Brett C. Luthans e Kyle W. Luthans (2015). *Organizational Behavior: An Evidence-Based Approach, 13th Ed.* 13 edition. Charlotte, North Carolina: Information Age Publishing. 536 pp. isbn: 978-1-68123-119-8.
- Maas, Andrew L et al. (2011). «Learning Word Vectors for Sentiment Analysis». Em: p. 9.
- Marreiros, Goreti, Carlos Ramos e José Neves (2005). «Emotion and Group Decision Making in Artificial Intelligence». Em:
- Marreiros, Goreti, Ricardo Santos et al. (2010). «Context-Aware Emotion-Based Model for Group Decision Making». Em: *IEEE Intelligent Systems* 25.2, pp. 31–39. issn: 1541-1672. doi: 10.1109/MIS.2010.46. url: <http://ieeexplore.ieee.org/document/5456410/> (acedido em 11/02/2020).
- Martinovski, Bilyana e Wenji Mao (2009). *Emotion as an Argumentation Engine: Modeling the Role of Emotion in Negotiation*. url: https://www.researchgate.net/publication/225999564_Emotion_as_an_Argumentation_Engine_Modeling_the_Role_of_Emotion_in_Negotiation (acedido em 16/06/2020).
- Moon, Henry et al. (2003). «Group Decision Process and Incrementalism in Organizational Decision Making». Em: *Organizational Behavior and Human Decision Processes* 92.1, pp. 67–79. issn: 0749-5978. doi: 10.1016/S0749-5978(03)00079-7. url: <http://www.sciencedirect.com/science/article/pii/S0749597803000797> (acedido em 06/02/2020).
- Ohana, Bruno e Brendan Tierney (2009). «Sentiment Classification of Reviews Using SentiWordNet». Em: *The IT&T 9th International Conference on Information Technology and Telecommunication 2009*. Dublin Institute of Technology, pp. 19–26.
- Oliver P., John e Srivastava Sanjay (1999). *Handbook of Personality, Second Edition: Theory and Research - Google Livros*. url: https://books.google.pt/books?hl=pt-PT&lr=&id=iXMqQ7wg-qkC&oi=fnd&pg=PA102&dq=The+Big+Five+trait+taxonomy:+History,+measurement,+and+theoretical+perspectives&ots=uEdba8iElb&sig=m0j3CRU9gq0C8cdYsNuiGkCKkSg&redir_esc=y#v=onepage&q=The%20Big%20Five%20trait%20taxonomy%3A%20History%2C%20measurement%2C%20and%20theoretical%20perspectives&f=false (acedido em 24/01/2020).
- Ortony, Andrew, Gerald L. Clore e Allan Collins (1990). *The Cognitive Structure of Emotions*. Cambridge University Press. 228 pp. isbn: 978-0-521-38664-7. Google Books: dA3JEEAp6TsC.
- Palomares, Iván, Luis Martínez e Francisco Herrera (2014). «A Consensus Model to Detect and Manage Noncooperative Behaviors in Large-Scale Group Decision Making». Em: *IEEE*

- Transactions on Fuzzy Systems* 22.3, pp. 516–530. issn: 1941-0034. doi: 10.1109/TFUZZ.2013.2262769.
- Pang, Bo e Lillian Lee (2004). «A Sentimental Education: Sentiment Analysis Using Subjectivity Summarization Based on Minimum Cuts». Em: *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*. ACL '04. Barcelona, Spain: Association for Computational Linguistics, 271–es. doi: 10.3115/1218955.1218990. url: <https://doi.org/10.3115/1218955.1218990> (acedido em 12/02/2020).
- Patodkar, Vaibhavi N e Sheikh I.R (2016). «Twitter as a Corpus for Sentiment Analysis and Opinion Mining». Em: *IJARCCCE* 5.12, pp. 320–322. issn: 22781021. doi: 10.17148/IJARCCCE.2016.51274. url: <http://ijarcce.com/upload/2016/december-16/IJARCCCE%2074.pdf> (acedido em 30/12/2019).
- Picard, Rosalind W. (2000). *Affective Computing*. MIT Press. 308 pp. isbn: 978-0-262-66115-7. Google Books: GaVncRTcb1gC.
- Plutchik, Robert (2001). «The Nature of Emotions: Human Emotions Have Deep Evolutionary Roots, a Fact That May Explain Their Complexity and Provide Tools for Clinical Practice». Em: *American Scientist* 89.4, pp. 344–350. issn: 0003-0996. JSTOR: 27857503.
- Roseman, Ira J, Martin S Spindel e Paul E Jose (1990). «Appraisals of Emotion-Eliciting Events: Testing a Theory of Discrete Emotions». Em: p. 17.
- Saaty, Thomas L. (2008). «Decision Making with the Analytic Hierarchy Process». Em: *International Journal of Services Sciences* 1.1, p. 83. issn: 1753-1446, 1753-1454. doi: 10.1504/IJSSCI.2008.017590. url: <http://www.inderscience.com/link.php?id=17590> (acedido em 06/02/2020).
- Saaty, Thomas L. e Kirti Peniwati (2013). *Group Decision Making: Drawing Out and Reconciling Differences*. RWS Publications. 484 pp. isbn: 978-1-888603-22-4. Google Books: rGwXAgAAQBAJ.
- Santos, Ricardo et al. (2011). «Personality, Emotion, and Mood in Agent-Based Group Decision Making». Em: *IEEE Intelligent Systems* 26.6, pp. 58–66. issn: 1941-1294. doi: 10.1109/MIS.2011.92.
- Serbedzija, Nikola B. e Stephen H. Fairclough (2009). «Biocybernetic Loop: From Awareness to Evolution». Em: *2009 IEEE Congress on Evolutionary Computation*. 2009 IEEE Congress on Evolutionary Computation (CEC). Trondheim, Norway: IEEE, pp. 2063–2069. isbn: 978-1-4244-2958-5 978-1-4244-2959-2. doi: 10.1109/CEC.2009.4983195. url: <http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=4983195> (acedido em 18/01/2020).
- Steunebrink, Bas et al. (2009). «The OCC Model Revisited». Em:
- Sutton, Charles (2012). «An Introduction to Conditional Random Fields». Em: *Foundations and Trends® in Machine Learning* 4.4, pp. 267–373. issn: 1935-8237, 1935-8245. doi: 10.1561/22000000013. url: <http://www.nowpublishers.com/article/Details/MAL-013> (acedido em 02/10/2020).
- Tsai, Jason et al. (2020). *A Study of Emotional Contagion with Virtual Characters* | SpringerLink. url: https://link.springer.com/chapter/10.1007/978-3-642-33197-8_8 (acedido em 16/06/2020).
- Van Kleef, Gerben A., Carsten K. W. De Dreu e Antony S. R. Manstead (2006). «Supplication and Appeasement in Conflict and Negotiation: The Interpersonal Effects of Disappointment, Worry, Guilt, and Regret». Em: *Journal of Personality and Social Psychology* 91.1, pp. 124–142. issn: 1939-1315(Electronic),0022-3514(Print). doi: 10.1037/0022-3514.91.1.124.

- Van Kleef, Gerben A., Carsten K. W. De Dreu e Antony S. R. Manstead (2004). «The Interpersonal Effects of Anger and Happiness in Negotiations.» Em: *Journal of Personality and Social Psychology* 86.1, pp. 57–76. issn: 1939-1315, 0022-3514. doi: 10.1037/0022-3514.86.1.57. url: <http://doi.apa.org/getdoi.cfm?doi=10.1037/0022-3514.86.1.57> (acedido em 15/07/2020).
- Van Kleef, Gerben, Carsten De Dreu e Antony Manstead (2010). «An Interpersonal Approach to Emotion in Social Decision Making: The Emotions as Social Information Model». Em: *Advances in Experimental Social Psychology* 42, pp. 45–96. doi: 10.1016/S0065-2601(10)42002-X.
- Walters, David e Geoff Lancaster (2000). «Implementing Value Strategy through the Value Chain». Em: *Management Decision* 38.3, pp. 160–178. issn: 0025-1747. doi: 10.1108/EUM0000000005344. url: <https://doi.org/10.1108/EUM0000000005344> (acedido em 10/01/2020).
- Woodall, Tony (2003). «Conceptualising 'Value for the Customer': An Attributional, Structural and Dispositional Analysis». Em: p. 45.